

U.S. Patent and Trademark Office
OFFICE OF THE CHIEF ECONOMIST
Economic Working Paper Series

Software Piracy and IP Management Practices: Strategic Responses to Product-Market Imitation

Wendy A. Bradley, *Assistant Professor*
Julian Kolev, *Economist*

USPTO Economic Working Paper No. 2021-3
August 2021

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Office of the Chief Economist or the U. S. Patent and Trademark Office. USPTO Economic Working Papers are preliminary research being shared in a timely manner with the public in order to stimulate discussion, scholarly debate, and critical comment. For more information about the USPTO's Office of the Chief Economist, visit www.uspto.gov/economics.



Software Piracy and IP Management Practices: Strategic Responses to Product-Market Imitation

Wendy A. Bradley
SMU Cox School of Business

Julian Kolev
USPTO Office of the Chief Economist

ABSTRACT: How do firms' IP strategies respond to sudden increases in product-market imitation? Using a 2001 technological shock that enabled rising software piracy, we implement an instrumental-variables estimator to compare a treatment group of at-risk-of-piracy firms with matched not-at-risk control firms. We find that rising piracy increases subsequent R&D spending, copyrights, trademarks, and patents for large, incumbent software firms. Furthermore, copyright and trademark filings precede those of patents, and firms with large patent portfolios disproportionately increase copyrights and trademarks following the shock. We conclude that piracy and similar competitive shocks push firms to innovate to stay ahead of imitator products, and that this effect is moderated by their existing patent portfolios. Our findings have implications for managers seeking to capture value from IP in knowledge-based industries.

JEL: O32, O34, D22, M21, L86

Keywords: piracy, software, innovation, intellectual property

ACKNOWLEDGEMENTS: Wendy Bradley: wabradley@smu.edu; Julian Kolev: julian.kolev@uspto.gov. We thank Scott Stern, Hong Luo, Ariel Stern, Pierre Azoulay, Danielle Li, and seminar participants at MIT, the 2019 Roundtable on Empirical Methods in IP, the 2019 NBER Summer Institute, the 2019 AOM Annual Meeting, the 2019 SMS Annual Meeting, and the 2020 NBER Productivity Seminar for their helpful comments. We gratefully acknowledge research assistance from Julian LaNeve and Addie Olson, and data consulting from Mike Barron.

DISCLAIMER: The views expressed are those of the individual authors and do not necessarily reflect official positions of the Office of the Chief Economist or the U. S. Patent and Trademark Office. USPTO Economic Working Papers are preliminary research being shared in a timely manner with the public in order to stimulate discussion, scholarly debate, and critical comment. For more information about the USPTO's Office of the Chief Economist, visit www.uspto.gov/economics.

1. INTRODUCTION

Large portfolios of intangible assets, such as patents, copyrights, and trademarks, have become the primary locus of value for publicly traded corporations (Berman 2019). As a result, the management of intellectual property (IP) has become central to firm strategy. Current literature focuses primarily on patenting, despite persistent survey evidence that patents remain the least-preferred mode of appropriability among managers, particularly in technology industries; instead, managers report using a combination of appropriation methods simultaneously (Arundel and Kabla 1998, W. Cohen et al. 2000, Hall et al. 2014, Levin et al. 1987). For example, an executive at Nokia once declared the company's most valuable asset to be a combination of patents, design rights, and trademarks protecting the underlying architecture of its cellular phones (Reitzig 2004).

Overlapping legal protections are central to a firm's ability to appropriate rents from its innovations and prevent imitation; without these, firms have little incentive to innovate (Teece 1986). In recent years, the threat of reverse-engineering and imitation has increased both the level of competition among firms (Roberts 1999), the potential for market failure in knowledge-based industries (Varian 2003), and the need for competitive repositioning (Wang and Shaver 2014). These challenges to appropriability are particularly salient for information goods. Copyrighted written works, lines of computer code, and trademarked words and images can often be reproduced and distributed rapidly at negligible marginal costs. This threat is amplified in digital goods, which have seen a dramatic reduction in the costs to reproduction and distribution, both legal and illegal, within and across national borders over the past twenty-five years (Goldfarb, Greenstein and C. E. Tucker 2015, Greenstein et al. 2013, Peitz and Waelbroeck 2006).¹ Shocks to the value of IP, whether by policy change or digital piracy, occur quite frequently and have the potential to

¹ Continuing the example of Nokia: the underlying proprietary software for Nokia's N-Gage video game console was hacked, reproduced, and distributed online quickly after launch (Charny 2003). In less than two years, Nokia discontinued producing games for the N-Gage console.

undermine firms' isolating mechanisms and strategies for preventing imitation through formal IP ownership (Hall 2004, Kortum and Lerner 1999).

Nevertheless, there is limited empirical investigation about how firms in knowledge-based industries manage their IP portfolio and few systematic studies about the strategic responses of firms to changes in their IP environment.² Current literature tends to focus on the impact of a shock to one form of IP; for example, the value of patents for software changed in the early 1990s (Campbell-Kelly 2004, Graham and Mowery 2006, Hall and MacGarvie 2010) and in 2014 with the *Alice v. CLS Bank* decision, while the value of copyrights for software changed in 1980 (Samuelson 1984, 2010). However, a comprehensive and dynamic understanding of firms' IP management practices, as part of firms' broader IP strategy to achieve competitive differentiation,³ requires a single policy or environmental shock that weakens (or strengthens) all forms of IP simultaneously. Only with such a shock is it possible to address the yet unanswered research question: how do large, public firms in knowledge-based industries alter their IP strategies when faced with a sudden rise in the threat of product-market imitation?

Our paper seeks to fill this gap in the literature with a two-stage approach. First, we leverage a sudden shock to firms' IP environment—digital piracy—that affects the value of all modes of formal IP equally as a natural experiment. Second, we build a dataset that matches the financial information of publicly traded software firms with (1) R&D expenditures and patent, copyright, and trademark counts, and (2) a unique dataset of pirated software. We leverage observable differences in the threat of piracy to treatment and control firms in an instrumental variable analysis using a novel instrument: mentions of piracy in firms' SEC-mandated annual 10-K filings. We compare changes in firms' IP strategies pre-piracy versus

² The question of firms' strategic responses to changes in their IP environment is related to but distinct from the extensive literature about the potential for complementarity versus substitutability of formal and informal IP (Amaral et al. 2008, Friedman et al. 1991, Hall et al. 2014, Horstmann et al. 1985, Teece 1986).

³ We define IP strategy as those aspects of a firm's broader strategy to obtain a competitive advantage through the ownership of formal IP rights. IP management practices (defined in Section 3.3) are observable decisions by firms in terms of IP input (R&D) and output (patents, copyrights, and trademarks), including their respective balance.

post-piracy across a sample of 106 publicly traded companies that generated over 40% of all revenue in the software industry from 1991-2000. The setting of software is ideal for this study: software products are in the unique position of being protected by patents, trademarks, and copyrights; they benefit disproportionately from cumulative innovation; and they exist in an economically important industry where firm reputation and network effects play major roles.

When comparing the IP strategies of software firms at risk of piracy (the treatment group) against those of not-at-risk firms (the control group), we find that our treatment group significantly increases its innovative activity after the piracy shock in terms of R&D expenditures and granted copyright, trademark, and patent applications. Our analysis also reveals a dynamic response: firms tend to increase their R&D expenditures and copyright filings sharply in the first two years following the piracy shock, while the impact on patents is most significant over longer horizons of three to seven years. Our results imply companies are quick to develop new products post-imitation via piracy and seek modes of appropriability with lead-time advantages, both informal ones via R&D and formal methods via copyrights. Trademarks have both short-term and long-term strategic uses and are positive and significant across both time horizons. Exploring further, we decompose the heterogeneous impact of piracy within our treatment group and find that firms with greater pro-patent management practices (i.e. firms that rely more heavily on patenting) tend to diversify their IP portfolio following the piracy shock. In effect, firms with more patents in the pre-period have significantly larger increases in copyrights and trademarks in the post-period, while the impact on R&D spending and patenting is relatively constant across our treatment group. These findings are consistent with prior evidence of complementarities among multiple methods of IP (Amara et al. 2008, Arora and Ceccagnoli 2006), suggesting that patent-rich firms aim to increase the value of their existing patent portfolios through alternative, complementary IP instruments.

This paper makes several contributions to the strategic management literature. First, we contribute to theory on the management of innovation (Grindley and Teece 1997; Rivette and Kline 2000; Somaya et

al 2007; Di Minin and Bianchi 2011) by explicating the special case of digital goods. Because such goods can be reproduced and distributed quickly and easily, IP protection methods and innovation practices with lead-time advantages, such as copyrights, may become increasingly important to managers (Edvardsson et al. 2006). Managers of software firms make use of a variety of formal IP instruments; yet, surprisingly, there is a lack of understanding about how firms handle the tradeoffs between the advantages and disadvantages of different forms of IP. Second, this paper contributes to discussions of the effectiveness of traditional methods for the appropriation of innovation in digital industries (Arora 1995, Goldfarb, Greenstein and C. Tucker 2015, Greenstein et al. 2013, Hall and Ziedonis 2001, Jaffe et al. 2001) and the impact of weak or absent IP rights on competition and innovation (Luo and Mortimer 2017, Mansfield 1986, Teece 1986, Zhao 2006). We extend this literature by jointly analyzing both inputs (R&D expenditures) and outputs (copyrights, trademarks, patents) to innovation within a single empirical framework. We provide the first large-sample analysis that systematically links IP management practices to heterogeneity in software firms' responses to their IP environment⁴ and break new ground by using the full spectrum of formal appropriability mechanisms to analyze firm strategy. Finally, this paper contributes to the broader literature on piracy and innovation (Ackermann, Bradley, and Cameron, 2020; Danaher, Smith and Telang 2014, Kariithi 2011, Peitz and Waelbroeck 2006, Yoon 2012), extending it to the software industry.

2. LITERATURE REVIEW

In this section, we provide background on studies related to software piracy, innovation, and the potential for complementarity among different IP instruments.

2.1 Digital Piracy in the Software Industry

Global information technology (IT) spending neared \$4 trillion in 2020, and the unauthorized reproduction and distribution of digital goods, or digital piracy, threatens the profitability of firms' digital business initiatives by providing identical, illegal substitutes for legitimate products (Costello and Rimol

⁴ The prior literature does include case studies of IP management practices (Bhaskarabhatla and Hegde 2014).

2021). The digital piracy of software, from productivity to enterprise software, differs critically from the digital piracy of products in other industries. First, when software are illegally copied, the copy is virtually indistinguishable from the original. This is unlike the piracy of digital goods in other markets, such as movies, music, or e-books, where the illegal copy may differ in quality or number of features (Wang, Li and Singh, 2018; Lu, Wang and Bendle, 2020; Ackermann, Bradley, and Cameron, 2020). For goods that can be reproduced with no distinguishable difference from the original, the theft of one version holds the potential to satisfy the demand of all interested consumers. This may be true for some physical goods, such as the photocopying of copyrighted journal articles (Besen 1986, Liebowitz 1985). However, unlike the theft of physical goods, the marginal cost of distributing digital products is near zero (Varian 2003). Furthermore, illegal copies of software can be made from other illegal copies of software, as opposed to a producer-made original as in the case of physical goods or digital media, in which a degradation of quality in subsequent copies occurs. This allows the producer to appropriate some of the consumer gains from illegal copies (Conner and Rumelt 1991).

Second, software relies heavily on network effects for product adoption and success. Digital piracy for goods in industries that generate network effects may lead to an increase in the critical mass of users necessary for increased product adoption, complementary product sales, and thus re-investment into R&D (Jaisingh, 2009; Conner and Rumelt, 1991; Shy and Thisse, 1999). In the software industry, the impact of network effects may lead a dominant firm to benefit from piracy if its product's popularity, authentic or otherwise, crowds out competitors, as seen with Linux's operating system and Apple's music platform (Goldfarb, Greenstein and C. Tucker 2015). Connor and Rumelt (1991) note the tradeoffs for software firms between asserting IP rights and exploiting the potential for positive network externalities when faced with the threat of digital piracy. As a result, the strategic responses and IP management practices of software firms facing rapid imitation via piracy may differ meaningfully from those of firms facing digital piracy in other markets.

2.2 Software Piracy and Innovation

Digital piracy has been a central concern of trade agreements (Awokuse and Yin 2010), intra-industry consortia (Jaisingh 2009), and IP litigation (Luo and Mortimer 2017). The debate on the economics of piracy in academic scholarship is well-informed and on-going (Athey and Stern 2015, Landes and Posner 1989). Much empirical research suggests that piracy reduces revenue and firm profit across industries, from software (Givon et al. 1995) to music (Bhattacharjee et al. 2007, Danaher, Smith, Telang, et al. 2014, Kretschmer and Peukert 2014, Rob and Waldfogel 2006) to movies (Ackermann, Bradley, and Cameron, 2020; De Vany and Walls, 2007; Bai and Waldfogel, 2012; Danaher and Smith, 2014; Ma *et al.*, 2014). The vast majority of theoretical literature predicts that piracy will decrease firm innovation as a result of lost revenues, which leads to lower firm profits and fewer incentives to innovate (Belleflamme and Peitz 2010). Although some argue that piracy may increase consumer awareness for a product, thus increasing revenues and incentives to innovate in certain industries, as in movies (Peukert et al. 2017) or music (Lee 2018), this mechanism may not be large enough to offset the revenue displacement effects of piracy (Lu et al. 2020).

Empirical evidence linking piracy and innovation, directly, remains mixed. In the movie industry, for example, Danaher and Smith (2017) show that piracy did not reduce the number of films made; rather, it changed the direction of filmmaking toward mainstream titles. Waldfogel and Aguiar (2018) find similar results, where piracy did not reduce the supply of new music. These studies find no evidence suggesting an impact of piracy on innovation. Empirical research on mobile applications offers some insight into the software setting by suggesting digital piracy has no discernable impact on major innovations, though it may disincentivize developers to refine or incrementally improve products (Miric and Jeppesen 2020). Generally, studies on piracy and innovation do not use IP or direct piracy data, and none investigate the relationship between innovation and piracy in the software industry.

2.3 Incentives to Innovate and Complementarity of IP

Competitive differentiation is achieved through “isolating mechanisms” (Rumelt 1984) that enable firms to strategically prevent the imitation of key resources or innovation. A firm’s ability to appropriate the

returns from its investments in innovations is a key driver of the firm's willingness to invest in innovative activity (Teece 1986). Much literature has focused on patents as the key appropriability mechanism for innovation within firms. This sustained interest may be explained by the growth of patenting (Hall 2004, Kortum and Lerner 1999) and the increasing value of patents over time (Bhaskarabhatla and Hegde 2014). However, patents are only one method of IP protection, and frequently neither the sole, nor most important indication of technological innovation for firms. Recent research suggests that trademarks are indicators of innovation in knowledge-intensive businesses (Amara et al. 2008, Gotsch and Hipp 2014, Mendonça et al. 2004) and that copyrights also track the introduction of new products and related innovations (Greenstein et al. 2013, Waldfogel 2012).

Overwhelming empirical evidence points to the fact that firms use a combination of appropriation mechanisms for innovation at the firm-level and even for the same individual invention (Hall et al. 2014). Seminal work by Levin, Klevorick, Nelson, and Winter (1987) was first among a series of innovation surveys to report that firm managers found patents less effective than alternative IP mechanisms, such as lead time, secrecy, and complementary sales and services; firms also reported using more than one appropriability method.⁵ In a survey by Paallysaho and Kuusisto (2008), firms in knowledge-based industries, including software consultancy and supply, preferred the use of trademarks and copyrights to patents.

Outside surveys, there is little empirical evidence measuring whether the return to using one type of IP increases the use or value of another form of IP. In one of these rare studies, Somaya and Graham (2006) show a complementary use of trademarks and copyrights by firms in the software industry. The authors analyze changes in litigation rates for copyrights and trademarks among firms in the software industry, controlling for firm size, age, R&D, income, IP management, and firm fixed effects, and they find that the residuals in the copyright and trademark litigation rates to be correlated. These results suggest that

⁵ See Hall, Helmers, Rogers, and Sena (2014) for a review of survey findings on manager preferences for methods of appropriability and the complementarity of formal and informal IP.

IP methods can be used as complements and pooled strategically, but the complementarity of IP depends strongly on market structure, competitive environment, and industry (Hall et al. 2014).

3. THEORY AND HYPOTHESES

In this section, we develop theoretical arguments for predicting the magnitude and sequence of innovative activity of firms following a shock to their IP environment. We focus on two separate outcomes. First, we consider the arguments put forth by previous literature regarding the impact of competitive shocks on expected innovative activity, including inputs to innovation (R&D) and outputs of innovation (patents, copyrights, and trademarks). Second, we introduce a new mechanism for this prediction that has received limited attention in previous work: pro-patent management practices. As a shock to competition, piracy may create incentives for firms to innovate, but this innovation will manifest according to a firms' IP management practices.

3.1 Impact of Piracy on Innovation

Within our context, we argue that piracy is best interpreted as a form of product-market competition similar to imitation or reverse engineering. Consider the following paragraph by Roberts (1999) about Sony Corporation:

After [Sony] introduces each new product, [it] experiences a rapid increase in sales and profits associated with that product. However, this leads other firms to reverse engineer the Sony product and introduce their own version. Increased competition leads the sales and profits associated with the new product to be reduced. Thus, at the level of individual products introduced by Sony, Sony apparently enjoys only very short-lived competitive advantages. (Roberts 1999: 657)

Reverse engineering is frequently how hackers illegally reproduce or pirate software code. In this way, piracy emulates the threat of increased competition described by Sony managers. Roberts (1999) continues:

However...Sony is able to constantly introduce new and exciting personal electronics products. No one of these products generate a sustained competitive advantage. However, over time, across several such product introductions, Sony's capability advantages do lead to a sustained competitive advantage." (Roberts 1999: 657)

This example echoes concepts about competition and innovation by Schumpeter (1939): firms may experience competitive pressures that provide them with incentives to differentiate as a market response (Aghion et al. 2005, Lahiri and Dey 2013, Posen and Martignoni 2018) or engage in rapid product innovation (D'Aveni and Gunther 1994, Roberts 1999, Schumpeter 1939) to increase firm performance and reduce the negative effects on revenues of imitation.

In our context of software, digital piracy may provide firms with incentives to innovate to improve their product offerings relative to competitors, legal or otherwise, to escape imitation and competition (Miric and Jeppesen 2020). This leads to the following hypothesis:

Hypothesis (H1). *Digital piracy leads firms to increase innovation in the form of R&D spending and granted patent, copyright, and trademark applications.*

3.2 Sequence of Innovation

Patents have been studied extensively in the context of appropriability, including in software (Bessen and Hunt 2007, Cockburn and MacGarvie 2011, Graham and Vishnubhakat 2013, Lemley and Cohen 2000). However, successful commercialization of a product also involves product differentiation, positioning, and advertising through trademarks (Bessen and Raskind 1991). The world's most valuable trademarks, valued at billions of dollars, are in the software industry (Stonefield 2011). In the U.S., computer programs are also protected as a form of written expression through copyright. Surprisingly, however, little is known about how software firms manage or combine the different forms of IP available to them.

Compared to patent applications, copyright and trademark applications are inexpensive, easily acquired, and lenient in their requirements, specifically with respect to originality. Patentable ideas take longer to conceive, design, and file, and they have longer lead-times to licensing and appropriation than do trademarks and copyrights. Firms are not required to file copyright registration to be legally protected; nevertheless, there are a variety of advantages to doing so.⁶ Copyright registrations are for expressions of ideas and not required to

⁶ See U.S. Copyright Office publication on *Copyright Basics*.

capture an entire software product. Instead, copyrighted lines of code may represent partial products. Similarly, trademark applications can precede product conceptualization and commercialization: firms may file “intent-to-use” registrations to be renewed every six months for up to three years without proof-of-concept. For these reasons, copyright and trademark applications often precede utility patent applications for the same product.

At the same time, the sequence of obtaining IP for innovations is not purely mechanical: software firms also seek IP strategically. Across innovation surveys of firms in both low- and high-tech industries, firms valued early product entry strategies and lead-time advantages over formal IP protection (Arundel 2001, Blind et al. 2006, W. Cohen et al. 2000, Harabi 1995, Hurmelinna-Laukkanen and Puumalainen 2007, Laursen and Salter 2006, Levin et al. 1987). Common explanations for why firms valued lead time over legal mechanisms of appropriability included the trade-off between disclosure of technical knowledge and the assurance of early protection through patents (Horstmann et al. 1985). Partial product disclosure through copyrights and trademarks allows firms to hide or bury complex technical knowledge about new products. Large, established firms, such as those in our study, have been shown to engage in this type of strategic IP behavior more often. For example, according to Arora (1997), large, established firms in 19th century Germany's high-tech sector minimized the disclosure of their inventions by combining patents and secrecy to deter entry. Similar to aggressive patenting behavior by firms, early investments in copyright and trademark applications may allow firms to extract rents from licensing sooner and avoid being “fenced in” by competition in markets for complex technology with overlapping IP rights (Ziedonis 2004). This leads to the following hypothesis:

Hypothesis (H2). *Digital piracy leads firms to increase R&D spending and granted copyright and trademark applications before increasing rates of patenting.*

3.3 Heterogeneity and Pro-Patent Management

Intangible assets play an increasingly important role for large, public software firms. In the late 1990s, half of the total market capitalization of S&P 500 companies came from intangible assets, including patents, copyrights, and trademarks. In 2019, 84 percent of the S&P 500's value comes from intangible

assets, where six of the top ten companies ranked on intangible value were in the software industry and all but two were in the technology sector (Berman 2019). Nevertheless, not all large, public software firms own patents.⁷ Despite overwhelming consensus on the importance of IP to firm strategy, there is little agreement on explanations for within-industry heterogeneity in IP ownership.

In a case study of IBM, Bhaskarabhatla and Hegde (2014) propose that within-industry heterogeneity in patenting is attributable to differences in IP management practices, and these practices are what led IBM to rebound in terms of profitability when faced with a change in the value of patents. IP management practices are “firm-specific practices that seek to profit from the organization’s intellectual assets,” as defined by Grindley and Teece (1997) and Bhaskarabhatla and Hegde (2014). Pro-patent management firms, or firms that rely heavily on patenting as a form of IP protection, may view patent ownership and licensing as a key part of the firm’s strategy execution and differ meaningfully from firms without pro-patent management practices in how they react to changes in competition and their IP environment.

In light of the prior research cited showing that patents and other modes of IP act as complements, pro-patent management firms, which, by definition, have a large cache of patents, will seek to maximize the value of these assets by filing for copyrights and trademarks in the wake of a competitive shock; thus, piracy, as a form of sudden product imitation, may motivate firms with pro-patent management policies to file for copyrights and trademarks at greater rates than firms with no or low pro-patent management practices. By diversifying their IP instruments, pro-patent management firms can strengthen the value of their current patents quickly and efficiently, as well as potential future patents. Outputs of innovation, including these formal IP instruments, require inputs to innovation, or R&D investment. This leads to the following hypothesis:

⁷ Figure A1 in the appendix illustrates the wide range of IP portfolios present in our treatment group prior to the piracy shock. Using a simplex plot, we highlight prominent firms that favor each of the three IP instruments we analyze.

Hypothesis (H3). *Digital piracy leads firms with greater pro-patent management practices to diversify their IP portfolios with increased R&D spending and granted copyright and trademark applications compared to firms with lower pro-patent management practices.*

4. SETTING, METHODS, AND DATA

The fundamental goal of our analysis is to document the strategic responses of firms facing a shock to their IP environment that impacts all forms of IP. Specifically, we focus on shocks that undermine a firm's competitive differentiation created through isolating mechanisms related to the prevention of imitation by competitors (Rumelt 1984); examples can include the introduction of generic pharmaceuticals, "copycat" products like those faced by Sony, or pirated versions of digital goods such as music, movies, or software. The ideal setting to examine the effect of such a shock on firms' IP strategies would be to identify two distinct but virtually identical firms (in non-overlapping customer markets), then randomly assign one of the firms to be subjected to a shock of increased product-market imitation. Importantly, to effectively analyze the impact of the shock on firms' IP portfolios, it would need to target all forms of the treated firm's IP, including not only formal IP rights, such as copyrights, trademarks, and patents, but also undisclosed IP such as trade secrets. Under this ideal experiment, we would measure the difference in subsequent R&D spending and applications for copyrights, trademarks, and patents by both the treatment and control firms to identify the causal impact of the competitive shock on the treated firm's IP strategy. To expand this analysis to the exploration of heterogeneous treatment effects, we would need multiple pairs of treatment and control firms with different pre-shock IP strategies.

While the ideal experiment described above is impossible to achieve under real-world conditions, we argue that a close approximation is possible by combining firm-level matching with an exogenous shock in the form of the 2001 release of a file-sharing technology: the BitTorrent communication protocol. In combination, these aspects of our empirical analysis allow us to identify similar firms based on pre-period characteristics and observe their differential response to the shock of BitTorrent-based digital piracy.

Moreover, by tracking a sample of firms with a wide range of pre-period IP strategies, we are able to estimate heterogeneous treatment effects in response to the shock of imitation through digital piracy.

4.1. Empirical Setting: The Release of BitTorrent as a Natural Experiment

Using the release of BitTorrent as a natural experiment has several advantages.⁸ First, the release date of BitTorrent on July 2, 2001, was sudden, arbitrary, and unexpected. Second, the release of BitTorrent uniquely affects large data files, such as software, whereas previous digital piracy technology, such as Napster, was reserved for small data files, such as music files or documents. Third, independent websites emerged quickly after BitTorrent's release and persist today. IP rights holders are unable to shut down BitTorrent due to the details of its technical architecture and corporate structure (Carrier 2010).⁹ This resulted in uninterrupted software piracy through the end of our sample period in 2007 and beyond.¹⁰ Major software firms did not adopt subscription-based offerings for their products until around 2013 with the introduction of Microsoft Office 365,¹¹ so these effects are principally outside our sample period. Thus, our quasi-experimental empirical design using BitTorrent offers the most direct way to identify the magnitude and direction of changes in software firms' IP strategies in response to the competitive shock of digital piracy.

4.2. Methods

⁸ For details regarding the exogeneity of the appearance of BitTorrent, see the "Empirical Setting" in the Appendix.

⁹ "The open and decentralized architecture of both the technology and the community meant that it was essentially impossible to shut down BitTorrent file-sharing... the ecosystem has remained impervious and still operates today much as it did over a decade ago," (Morris 2018). BitTorrent's chief indexing website, The Pirate Bay (TPB), was also impervious to legal action. TPB founders explain: "If the police decide to raid us again there are no servers to take, just a transit router. If they follow the trail to the next country and find the load balancer, there is just a disk-less server there. In case they find out where the cloud provider is, all they can get are encrypted disk-images. They have to be quick about it too, if the servers have been out of communication with the load balancer for 8 hours they automatically shut down."

¹⁰ We end our sample in 2007 to avoid the confounding effect of the 2008 financial crisis. In robustness tests, we extend our analysis through 2013. Beyond this point, our analysis would be influenced by the 2014 Supreme Court decision in *Alice Corp. v. CLS Bank International*, declaring that the implementation of abstract claims on a computer is not patentable subject matter. This decision is viewed as greatly limiting the potential patentability of software.

¹¹ Microsoft (MS) Office 365 was released on February 27, 2013 as a subscription-based software in tandem with a non-subscription-based edition. MS Office 2019 was the first version to require a subscription to access the full features of the software suite.

Our empirical approach is based on a combination of a difference-in-differences framework, coarsened exact matching, and an instrumental-variables design around the software piracy event described in the previous section. Below, we describe these methods in turn.

4.2.1 Difference-in-Differences Framework

We investigate the impact of software piracy on the innovative activity of *at-risk* software firms—the treatment sample—before and after the release of the BitTorrent protocol in 2001, as compared to *not-at-risk* software firms—our control sample. We therefore combine this difference in *piracy risk* with a difference across time to generate a difference-in-differences estimation framework. We measure piracy risk using pre-period 10-K filings that include the exact strings of “pirate” or “piracy.” This approach has a number of important benefits: annual 10-K filings are mandated by the Securities and Exchange Commission (SEC) for all firms with at least \$10M in assets and more than 2,000 equity-holders. This report is required to include a discussion of risk factors, and firms are prohibited from making false or misleading statements, and from omitting material information in their 10-K filings. Because of this, we argue that mentions of piracy in a firm’s 10-K filings are an effective way to capture at-risk-of-piracy firms within our sample. Firms that do not mention pirates or piracy in any 10-K prior to the treatment date are used as the control group.¹²

4.2.2 Coarsened Exact Matching

As with any difference-in-differences methodology, we require our control group to be a close approximation to our treatment group prior to the shock of the BitTorrent protocol’s release. To ensure that our sample meets this requirement, we implement coarsened exact matching (CEM) (Iacus et al. 2012) at the firm level based on pre-period characteristics. The natural tradeoff of this approach is that requiring an exact match along too many dimensions, even after coarsening, would eliminate a growing fraction of firms

¹² Most firms in our control group produce software designed for businesses, such as network infrastructure and hosting applications, document storage, or payroll and accounting products. Business customers are less likely to engage in piracy because they may deduct legal purchases of software as a business expense, require post-purchase services, or must file audited financial statements detailing their assets and cash flows. As a result, the firms in our control group are less likely to change their innovative activity and overall IP strategy based on a rising threat of piracy.

from our estimation sample. We therefore match only on the variables of firm age and R&D expenditures, both log-transformed and measured at the end of the pre-period. Importantly, as seen in Table 2, this matching process balances our sample across all pre-period characteristics that we track, and retains all but one of the firms in our treatment group.¹³

4.2.3 Instrumental Variable Strategy

While the combination of matching and a difference-in-differences framework is sufficient to identify the directional impact of piracy on innovation, it falls short in two important areas. First, it does not offer an estimate of the effect size of realized piracy, and second, it does not fully address concerns about the potential for endogeneity. Endogeneity may exist along a number of dimensions in our setting. Perhaps the most problematic of these is the possibility that pirates are more likely to target the software of more innovative firms. Specifically, this would generate a positive correlation between our measure of piracy and the unobserved “innovativeness” of the firms in our sample, and generate an upward bias in standard OLS and Poisson estimates. To achieve an accurate estimate of the effect size of piracy’s impact on innovative activity, we implement an instrumental-variables estimation strategy based on our measure of firms’ piracy-mentioning 10-K filings in the pre-period. In addition to the already-discussed benefits of mandated annual disclosure and the prohibition on misleading statements and omissions of material information, this measure is also likely to satisfy the necessary exclusion restriction: mentions of piracy in firms’ 10-K filings during the pre-period of 1991-2000 are not likely to influence firms’ innovative activities through channels other than piracy in the post-period after controlling for observable firm characteristics. This exclusion restriction becomes even stronger when combined with our difference-in-differences framework: we instrument for observed levels of piracy using our continuous measure of piracy-mentioning 10-K filings interacted with a range of time periods to track either the overall impact or the dynamic effects of the piracy

¹³ Note that we implement CEM weights throughout our regression analysis in order to properly balance our sample. See Iacus et al (2012) for the full details of the CEM methodology.

shock across our outcomes of interest.¹⁴ If the exclusion restriction is satisfied, our estimates will reflect the local average treatment effect of piracy rates on the innovative activity of the firms in our sample.

4.3 Data Sources and Key Variables

We rely on four interlinked data sources. First, focusing on publicly traded firms, we collect financial information from COMPUSTAT's Fundamentals Annual. Second, we obtain copyright data from the U.S. Copyright Office, and third, we collect patent and trademark data from the U.S. Patent and Trademark Office.¹⁵ Finally, we track pirated software from The Pirate Bay (TPB).

Our empirical analysis focuses on a matched sample consisting of a treatment group of at-risk software firms and a control group of not-at-risk software firms. To construct our sample of firms, we first identified all firms in COMPUSTAT during our pre-period of 1991-2000 that fell within NAICS code 511210 or SIC code 7372 (both indicating the software industry), and that reported annual revenue greater than or equal to \$100M at least once. We exclude video game and media-focused firms, subsidiaries, and investment-holding companies, as well as firms that exit the sample prior to 2000. Finally, we implement the CEM methodology described above to balance our treatment and control groups. As shown in Table 2, this yields well-balanced treatment and control groups across not only the dimensions of matching, but also over a wide range of measures of financial and innovative activity. Our final sample consists of 41 at-risk software firms and 65 not-at-risk software firms. Together, our sample accounts for over 40% of all reported pre-period revenue by publicly traded firms in the software industry.

Our outcome variables cover a range of firms' innovative activities, beginning with the input of R&D expenditures captured in our focal firms' annual income statements. For patents, we focus on the application dates of granted utility patents assigned to our focal firms and discard reissued patents. We also track citation-weighted patent counts, tracking citations that accrue within ten years of the grant date of each

¹⁴ This methodology closely follows the approach of Bertrand et al (2019).

¹⁵ Specifically, we use the Trademark Case Files dataset (www.uspto.gov/ip-policy/economic-research/research-datasets/trademark-case-files-dataset) and the PatentsView dataset (www.patentsview.org), both from the USPTO.

patent and excluding self-citations. R&D expenditures and patents can be interpreted as traditional measures of innovation input and output, respectively. However, to offer a comprehensive analysis of firms' innovative activity and intellectual property, we also hand-collect copyright and trademark filings, which we interpret as reflecting new features and products introduced by our focal firms. We aggregate these measures at the firm-year level.

To explain the above outcome measures, we incorporate a broad range of independent variables. For our instrumental-variables analysis, we would need access to direct observations of firm-level piracy to accurately estimate the effect size of its impact on innovation. Due to its illegal nature, piracy is not tracked by traditional research datasets, and until recently previous analyses have generally had to estimate piracy without observing it directly (Business Software Alliance, 2011). To resolve this issue, we collect a novel dataset of all torrents uploaded from 2004-2013 via the BitTorrent protocol on TPB and use text-matching algorithms to match this data to our sample of firms. TPB was created in 2003 and remained the most-visited torrent directory on the Internet from 2003 until November 2014.¹⁶ Importantly, because the TPB did not exist during the first two years of the BitTorrent protocol, we use a time-invariant measure of realized piracy: the log-transformed total number of times each firm's products were posted on TPB during our sample period.¹⁷ To our knowledge, this is the first use of TPB data for both the software industry and the analysis of innovation and firm strategy, although two studies recently used TPB data for research on demand in the movie industry with similar approaches to identification (Lu, Wang and Bendle, 2020; Ackermann, Bradley and Cameron, 2020). As discussed above, we instrument for observed levels of piracy using the number piracy-mentioning 10-K filings (as well as the total number of 10-K filings) submitted to the SEC with respect to the years in our pre-period: 1991-2000.

¹⁶ The TPB's consistent popularity implies that the piracy patterns observed there are representative of worldwide trends. At its peak in 2008, TPB was the 98th most visited website across the worldwide web, according to Alexa data.

¹⁷ While this approach would introduce look-ahead bias under traditional OLS and Poisson specifications, it does not interfere with our instrumental-variables estimators.

Our remaining explanatory variables are constructed using income statement and balance sheet measures from COMPUSTAT, again focusing on pre-period characteristics. Finally, we implement calendar-year fixed effects in all specifications, and use firms' IPO dates to construct firm-age controls for our panel analysis. Tables 1 and 2 provide summary statistics and the details of sample balancing between treatment and control groups under our CEM methodology.

Table 1: Summary Statistics

Variable	N. of Obs.	Mean	Std. Dev.	Min	Max
Outcome Measures:					
R&D Expenditures (\$M)	1,165	106.44	460.79	0	7085.10
Copyright Count	1,165	11.67	62.90	0	1452
Trademark Count	1,165	4.70	13.96	0	197
Patent Count	1,165	23.13	183.45	0	3521
Citation-Weighted Patent Count	1,165	391.09	2598.80	0	40508
Firm-Year Covariates:					
Year	1,165	1999.96	4.07	1991	2007
Firm Age (from IPO year)	1,165	6.53	4.86	0	26
Firm-Level Covariates:					
Ever Pirated	106	0.17	0.38	0	1
Total Times Pirated	106	15.16	76.65	0	539
Pre-Period Piracy Indicator	106	0.39	0.49	0	1
Pre-Period Piracy 10-Ks	106	1.25	1.95	0	7
Pre-Period SEC 10-K Filings	106	4.19	1.58	1	8
IPO Year	106	1994.53	4.08	1981	2000
Log(Pre-Period R&D)	106	3.27	0.90	1.65	7.39
Log(Pre-Period Revenue)	106	4.92	0.94	3.66	9.24
Log(Pre-Period Market Value)	106	6.93	1.39	4.01	11.96
Log(Pre-Period Assets)	106	5.31	1.03	3.63	9.71
Log(Pre-Period Employees)	106	0.66	0.45	0.16	3.08
Log(Pre-Period Acquisitions)	106	1.31	1.13	0.00	6.57
Average Pre-Period Net Margin	106	-0.10	0.28	-1.25	0.28
Average Pre-Period Leverage	106	0.21	0.97	-4.66	5.07
Average Pre-Period Revenue Growth	106	0.32	0.25	-0.39	0.92
Log(Pre-Period Copyrights)	106	0.82	1.13	0.00	5.42
Log(Pre-Period Trademarks)	106	1.25	0.72	0.00	4.28
Log(Pre-Period Patents)	106	0.75	0.95	0.00	5.97

Table 2: Sample Balancing

This table compares pre-period characteristics between our treatment firms (SEC-piracy sub-sample) and control firms (non-SEC-piracy sub-sample). The unbalanced sample includes all firms and is equally-weighted. The balanced sample is weighted using coarsened exact matching based on the log of maximum pre-period R&D expenditures and the log of firm age in the year 2000.

Pre-Period Firm Characteristics	Treatment Firms: Unbalanced Sample		Control Firms: Unbalanced Sample		Difference in Means: Unbalanced Sample	
	Mean	SD	Mean	SD	Difference	P-Value
Log(Firm Age in 2000)	1.603	0.701	1.780	0.786	-0.177	(0.198)
Log(Max Pre-Period R&D)	4.186	1.080	3.762	1.220	0.424	(0.048)
IPO Year	1994.8	4.111	1993.1	6.131	1.662	(0.070)
Log(Pre-Period R&D)	3.455	1.035	3.109	1.166	0.346	(0.090)
Log(Pre-Period Revenue)	5.078	1.047	5.016	1.056	0.062	(0.752)
Log(Pre-Period Market Value)	7.416	1.516	6.627	1.504	0.789	(0.007)
Log(Pre-Period Assets)	5.568	1.170	5.272	1.091	0.296	(0.171)
Log(Pre-Period Employees)	0.726	0.477	0.747	0.592	-0.021	(0.828)
Log(Pre-Period Acquisitions)	1.221	1.027	1.333	1.206	-0.111	(0.586)
Average Pre-Period Net Margin	-0.112	0.310	-0.094	0.282	-0.018	(0.748)
Average Pre-Period Leverage	0.078	1.010	0.330	0.966	-0.252	(0.181)
Average Pre-Period Revenue Growth	0.363	0.260	0.249	0.273	0.115	(0.023)
Log(Pre-Period Patents)	1.031	1.191	0.633	0.863	0.164	(0.462)
Log(Pre-Period Trademarks)	1.403	0.786	1.138	0.727	0.265	(0.070)
Log(Pre-Period Copyrights)	0.932	1.262	0.768	1.004	0.398	(0.057)
Pre-Period SEC 10-K Filings	4.119	1.699	4.416	1.615	-0.297	(0.346)
Number of Firms	42		89			
Share of Pre-Period Industry Revenue	24.2%		38.2%			
Pre-Period Firm Characteristics	Treatment Firms: CEM Weighted Sample		Control Firms: CEM Weighted Sample		Difference in Means: CEM Weighted Sample	
	Mean	SD	Mean	SD	Difference	P-Value
Log(Firm Age in 2000)	1.602	0.710	1.615	0.740	-0.013	(0.940)
Log(Max Pre-Period R&D)	4.147	1.064	4.123	0.986	0.024	(0.917)
IPO Year	1994.8	4.158	1994.5	4.941	0.299	(0.803)
Log(Pre-Period R&D)	3.426	1.030	3.406	0.898	0.019	(0.929)
Log(Pre-Period Revenue)	5.056	1.049	5.037	0.994	0.018	(0.938)
Log(Pre-Period Market Value)	7.376	1.512	7.110	1.176	0.266	(0.356)
Log(Pre-Period Assets)	5.521	1.143	5.466	1.021	0.055	(0.819)
Log(Pre-Period Employees)	0.716	0.478	0.733	0.528	-0.017	(0.886)
Log(Pre-Period Acquisitions)	1.239	1.034	1.502	1.421	-0.264	(0.353)
Average Pre-Period Net Margin	-0.112	0.314	-0.107	0.275	-0.005	(0.941)
Average Pre-Period Leverage	0.075	1.022	0.180	0.582	-0.106	(0.536)
Average Pre-Period Revenue Growth	0.360	0.262	0.335	0.235	0.025	(0.634)
Log(Pre-Period Patents)	0.992	1.180	0.761	0.759	0.105	(0.702)
Log(Pre-Period Trademarks)	1.386	0.788	1.278	0.679	0.107	(0.513)
Log(Pre-Period Copyrights)	0.955	1.269	0.850	1.154	0.232	(0.283)
Pre-Period SEC 10-K Filings	4.098	1.715	4.085	1.549	0.013	(0.972)
Number of Firms	41		65			
Share of Pre-Period Industry Revenue	23.8%		17.8%			

4.4 Estimated Equations and Regression Specifications

Our analysis focuses on a difference-in-differences estimator for the impact of rising software piracy following the release of the BitTorrent protocol in 2001 and the founding of TPB in 2003 as the largest BitTorrent indexing website. Based on these dates, our first difference is captured by the “Post,” period, which is defined as an indicator equal to one from 2001 through 2007. In extended analysis, we also decompose the “Post” period into a “Window” period covering 2001 and 2002, and a “Post-Long” period covering 2003 through 2007 (inclusive).

Our second difference focuses on comparing at-risk-of-piracy and not-at-risk companies, using piracy-mentioning 10-K filings in the pre-period. However, at-risk status in the pre-period does not correlate perfectly with realized piracy: some firms in the treatment group remain piracy-free, even as some in the control group have their products pirated. Our empirical strategy, therefore, parallels an intent-to-treat analysis: to capture the treatment-on-treated effect size for the impact of piracy on innovative activity, we implement an instrumental-variables methodology (Angrist and Pischke 2014). Specifically, following Bertrand, Black, Jensen, and Lleras-Muney (2019), we use the interaction between pre-period at-risk status and the post-piracy time periods above to instrument for the impact of rising piracy on the firms in our sample. We capture the impact of piracy over time by instrumenting for the interaction between our measure of observed piracy, $\text{Log}(\widehat{\text{TimesPirated}})_i$ and the same post-piracy time periods that we use when constructing our instrumental variables. Our baseline analysis therefore estimates the following equation:

$$\begin{aligned} \text{InnovativeOutcome}_{it} &= \beta_0 X_{it} + \beta_1 \text{Log}(\widehat{\text{TimesPirated}})_i + \beta_2 \text{Post}_t \\ &\quad + \beta_{12} [\text{Log}(\widehat{\text{TimesPirated}})_i * \text{Post}_t] + \gamma_t + \epsilon_{it} \end{aligned}$$

In the above specification, X_{it} captures our control variables for firm i in year t , including the pre-period characteristics described in the previous section as well as log-transformed firm age and the total number of pre-period 10-K filings. All specifications also include calendar-year fixed effects γ_t , which absorb the

direct impact of the “Post” variable. We use log-linear 2SLS specifications for R&D expenditures, and IV-Poisson GMM specifications for our count variables.¹⁸ The primary coefficient of interest is β_{12} , which captures the local average treatment effect of piracy rates on firms’ innovative activity.

In addition to the baseline specification above, we also replicate our analysis while replacing the “Post” period with “Window” and “Post-Long” measures to track the dynamic responses of firms to the piracy shock. Finally, we introduce a triple-interaction with the number of pre-period patent filings to investigate the possibility of heterogeneous responses to the piracy shock within our treatment group. In both of these extensions, we add the corresponding interactions with our instrument in order to address the endogeneity of realized piracy.

5. RESULTS

The objective of this study is to provide a nuanced understanding of the ways in which large, public software firms respond to an exogenous increase in the threat of product-market imitation. We capture this effect by evaluating the impact of digital piracy on firms’ innovative activity. Our results begin with Figures 1 through 4, which capture the four primary channels through which we track firms’ IP strategies: R&D spending and patent, trademark, and copyright filings. These figures track the differences between our treatment and control firms over time; unlike our regression analysis, we use a binary measure of piracy risk (based on pre-period 10-K filings) to define our treatment and control groups. Moreover, these figures present the raw difference-in-means for each calendar year, without controlling for any additional covariates. While these methodological choices are helpful in presenting a clear depiction of the raw data in our sample, they also produce wide standard errors and lead the confidence intervals around our year-by-year point estimates to often include zero. Despite this, we find strong effects of approximately 0.5 for R&D expenditures and patents, and 1.0 or more for copyrights and trademarks. Since we use exponential models for all of our

¹⁸ Note that we do not include firm-level fixed effects in our primary specifications, because they cause IV-Poisson estimators to become inconsistent. In Table A6, we include reduced-form results with firm-level fixed effects.

outcome variables, these coefficients reflect increases of over 50% and 150%, respectively, in firms' innovative output following the piracy shock.

Figure 1: Treatment Effect on R&D Expenditures

Coefficients reflect a log-linear specification with standard errors clustered by firm

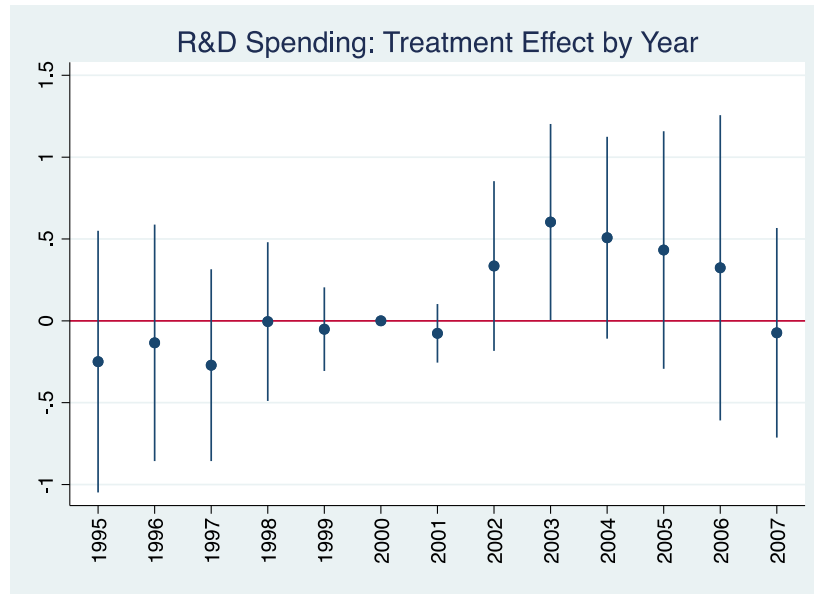


Figure 2: Treatment Effect on Copyrights

Coefficients reflect a Poisson specification with standard errors clustered by firm

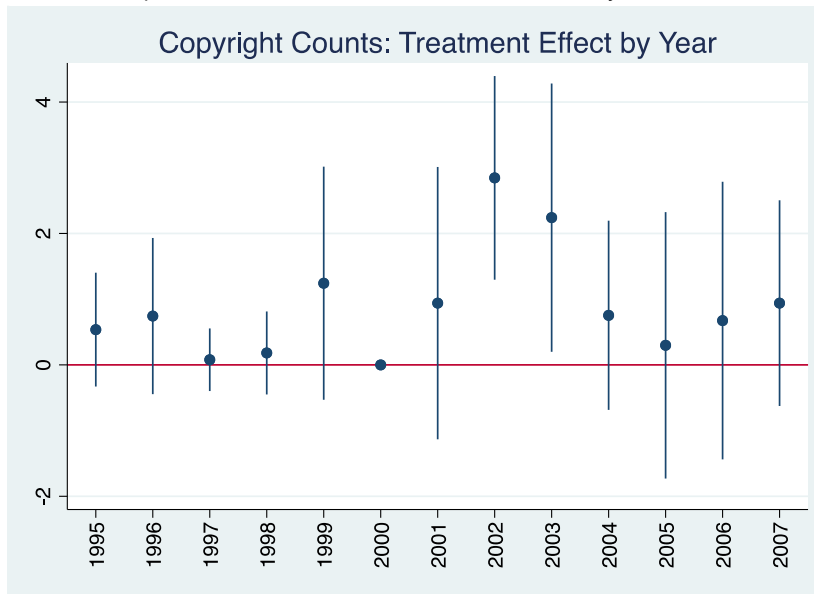


Figure 3: Treatment Effect on Trademarks

Coefficients reflect a Poisson specification with standard errors clustered by firm

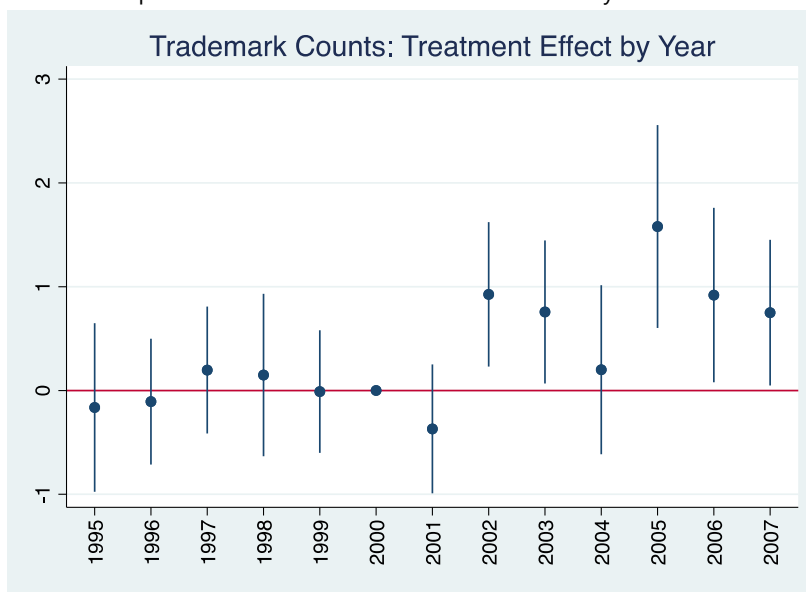
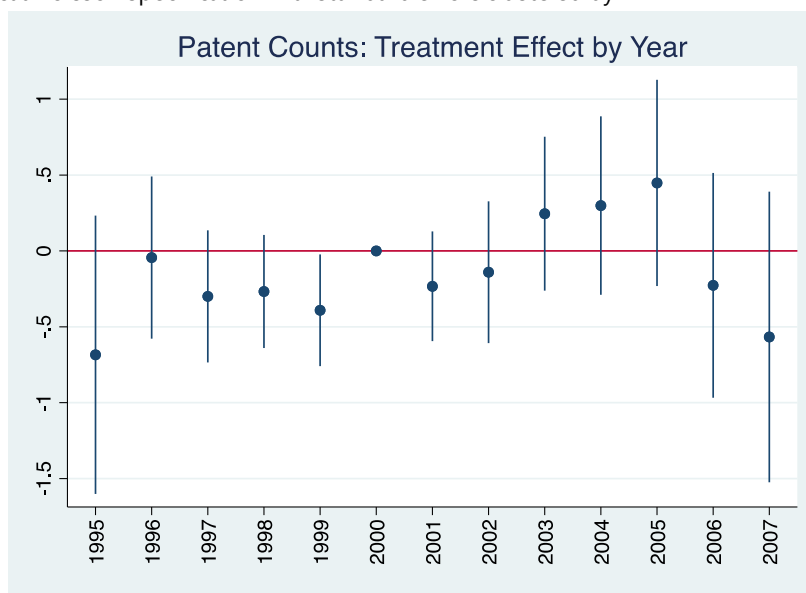


Figure 4: Treatment Effect on Patents

Coefficients reflect a Poisson specification with standard errors clustered by firm



5.1 Baseline Instrumental-Variables Estimates

As described in the previous section, the difference-in-differences analysis in our Figures can demonstrate the direction of impact for our piracy shock, but it does not offer an accurate estimate for the effect size of piracy's impact on innovative activity. We therefore complement the results in Figures 1

through 4 using an instrumental-variables (IV) analysis, with our results presented in Tables 3 and 4. Table 3 highlights the first stage of our two-stage analysis, demonstrating a highly robust positive relationship between pre-period at-risk status and the number of times firms' products are pirated on TPB in the post-period. Specification 3 of Table 3 is the model that carries over to Table 4, where we perform the second stage of the IV estimation. Unsurprisingly, we find the same pattern of results here as we did in our difference-in-differences analysis. However, the coefficients now reflect the elasticity of firms' innovative activity with respect to piracy rates. Specifically, the elasticity of 0.28 for R&D expenditures indicates that a ten-percent increase in the rate of piracy leads to an increase of 2.8% in R&D expenditures; for the average firm in our sample, for example, this would represent an increase of approximately \$3M per year. Similarly, we would expect increases of 2.0% in copyrights, 1.6% in trademarks, 0.8% in patents, and 0.5% in citation-weighted patents. As piracy rates were both increasing rapidly over time and highly concentrated in our treatment sample, these elasticities are consistent with the large increases in innovative activity presented in Figures 1 through 4.

Table 3: Instrumental Variables Stage 1 - Piracy 10-Ks and Observed Piracy

VARIABLES	(1) DV = Log(Total Times Pirated)	(2) OLS DV = Log(Total Times Pirated)	(3) DV = Log(Total Times Pirated)	(4) DV = Total Times Pirated	(5) Poisson DV = Total Times Pirated	(6) Poisson DV = Total Times Pirated
<u>Instrument:</u>						
Pre-Period Piracy 10-Ks	0.477 (0.000)	0.380 (0.000)	0.309 (0.000)	0.992 (0.000)	0.379 (0.002)	0.504 (0.000)
<u>Controls:</u>						
Pre-Period SEC 10-K Filings		0.046 (0.772)	-0.029 (0.856)		0.755 (0.165)	0.650 (0.119)
Log(Firm Age)		0.055 (0.656)	0.171 (0.091)		0.319 (0.003)	0.288 (0.005)
IPO Year		-0.080 (0.087)	-0.059 (0.253)		0.297 (0.137)	0.195 (0.255)
Log(Pre-Period R&D)		0.284 (0.247)	0.226 (0.336)		2.490 (0.051)	2.975 (0.028)
Log(Pre-Period Revenue)		0.352 (0.586)	0.278 (0.632)		1.939 (0.074)	1.195 (0.226)
Log(Pre-Period Market Value)		0.231 (0.270)	0.066 (0.728)		0.128 (0.780)	0.418 (0.369)
Log(Pre-Period Assets)		-0.138 (0.613)	-0.168 (0.479)		-0.617 (0.548)	0.338 (0.626)
Log(Pre-Period Employees)		-0.520 (0.454)	-0.674 (0.289)		-4.577 (0.000)	-5.474 (0.003)
Log(Pre-Period Acquisitions)		-0.275 (0.135)	-0.265 (0.137)		-0.383 (0.295)	-0.671 (0.118)
Average Pre-Period Net Margin		0.275 (0.683)	0.390 (0.515)		-1.111 (0.646)	-1.433 (0.358)
Average Pre-Period Leverage		0.156 (0.196)	0.162 (0.142)		0.350 (0.329)	0.885 (0.073)
Average Pre-Period Revenue Growth		0.033 (0.975)	0.761 (0.443)		-1.021 (0.666)	-3.484 (0.162)
Log(Pre-Period Copyrights)			0.262 (0.082)			0.061 (0.728)
Log(Pre-Period Trademarks)			0.237 (0.138)			-1.746 (0.017)
Log(Pre-Period Patents)			0.228 (0.193)			0.462 (0.135)
Year Fixed Effects	N	Y	Y	N	Y	Y
Observations	1,165	1,165	1,165	1,165	1,165	1,165
R-squared	0.414	0.577	0.622	0.771	0.892	0.900
F-test on Instruments	20.63	25.96	13.94			
Chi-square test on Instruments				19.41	9.84	15.66

P-values in parentheses, based on robust standard errors clustered by firm.

Table 4: Instrumental Variables Stage 2 - Impact of Piracy on Innovative Activity

VARIABLES	(1) 2SLS DV = Log(R&D)	(2) IV-Poisson DV = Copyright Count	(3) IV-Poisson DV = Trademark Count	(4) IV-Poisson DV = Patent Count	(5) IV-Poisson DV = Cite- Weighted Patent Count
<u>Instrumented Variables:</u>					
Log(Times Pirated) X Post (2001-2007)	0.280 (0.026)	0.197 (0.022)	0.157 (0.001)	0.078 (0.009)	0.048 (0.097)
Log(Total Times Pirated)	-0.126 (0.119)	-0.000 (0.999)	-0.084 (0.047)	0.055 (0.379)	0.170 (0.003)
<u>Selected Controls:</u>					
Pre-Period SEC 10-K Filings	-0.008 (0.856)	0.061 (0.752)	0.115 (0.108)	0.207 (0.161)	0.116 (0.457)
Log(Firm Age)	0.680 (0.000)	1.690 (0.001)	0.021 (0.895)	0.460 (0.292)	0.132 (0.605)
IPO Year	0.081 (0.009)	0.215 (0.014)	0.033 (0.267)	0.127 (0.019)	0.082 (0.131)
Log(Pre-Period R&D)	0.980 (0.000)	-0.084 (0.786)	-0.033 (0.803)	0.133 (0.655)	0.088 (0.707)
Log(Pre-Period Copyrights)	0.067 (0.211)	0.885 (0.000)	-0.013 (0.694)	-0.162 (0.023)	-0.292 (0.000)
Log(Pre-Period Trademarks)	0.044 (0.580)	0.509 (0.016)	1.320 (0.000)	0.219 (0.449)	-0.150 (0.544)
Log(Pre-Period Patents)	0.024 (0.616)	-0.159 (0.263)	-0.097 (0.009)	0.757 (0.000)	0.688 (0.000)
Additional Pre-Period Characteristics	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y
Observations	1,165	1,165	1,165	1,165	1,165
P-Value for Test of Endogeneity	0.504	0.772	0.103	0.409	0.590
N. of Firms	106	106	106	106	106

P-values in parentheses, based on robust standard errors clustered by firm.

Taking stock of the full spectrum of results in Figures 1 through 4 and in Table 4, we find strong support for the conclusion that the rise of piracy led to a significant increase in a wide range of innovative activity, from inputs like R&D expenditures to outputs such as copyrights, trademarks, and patents. We therefore confirm Hypothesis 1 across all of our outcome variables.

5.2 Dynamic Responses to Software Piracy

In addition to evaluating the overall impact of piracy on each of our outcome measures, our setting allows us to estimate the dynamics of firms' IP strategies. We present our findings in Table 5, where we differentiate between the "Window" (2001-2002) and "Post-Long" (2003-2007) time periods. We find that the increases in copyrights and trademarks are quite rapid following the release of the BitTorrent protocol, while the response in patent filings does not manifest until several years later. Unsurprisingly, R&D expenditures, which support all forms of innovation and IP, exhibit a strong response in both time periods. Also worth noting is the fact that all outcome measures except copyrights exhibit an acceleration as we move from the "Window" to the "Post-Long" period;¹⁹ this suggests that the effects we document are quite durable, lasting the better part of a decade following our piracy shock. Overall, these findings lend support to Hypothesis 2: copyrights and trademarks form the foundation of firms' short-term reactions to product-market imitation, while patents are used as a longer-term strategic response.

¹⁹ The fact that copyrights show a declining effect over time is a puzzle worthy of further investigation.

Table 5: Instrumental Variables Stage 2 - Dynamic Impact of Piracy

	(1) 2SLS	(2) IV-Poisson	(3) IV-Poisson	(4) IV-Poisson	(5) IV-Poisson DV = Cite- Weighted Patent Count
VARIABLES	DV = Log(R&D)	DV = Copyright Count	DV = Trademark Count	DV = Patent Count	
<u>Instrumented Variables:</u>					
Log(Times Pirated) X Window (2001-2002)	0.198 (0.054)	0.283 (0.005)	0.108 (0.054)	0.011 (0.746)	-0.009 (0.759)
Log(Times Pirated) X Post-Long (2003-2007)	0.317 (0.024)	0.137 (0.137)	0.194 (0.000)	0.112 (0.003)	0.089 (0.023)
Log(Total Times Pirated)	-0.126 (0.118)	0.004 (0.965)	-0.084 (0.048)	0.043 (0.496)	0.160 (0.004)
<u>Selected Controls:</u>					
Pre-Period SEC 10-K Filings	-0.008 (0.865)	0.064 (0.739)	0.115 (0.110)	0.201 (0.167)	0.114 (0.459)
Log(Firm Age)	0.689 (0.000)	1.614 (0.001)	0.054 (0.737)	0.670 (0.156)	0.278 (0.340)
IPO Year	0.081 (0.010)	0.209 (0.014)	0.038 (0.201)	0.146 (0.009)	0.097 (0.084)
Log(Pre-Period R&D)	0.979 (0.000)	-0.084 (0.778)	-0.036 (0.784)	0.129 (0.666)	0.082 (0.725)
Log(Pre-Period Copyrights)	0.066 (0.217)	0.884 (0.000)	-0.014 (0.685)	-0.154 (0.028)	-0.287 (0.000)
Log(Pre-Period Trademarks)	0.044 (0.577)	0.495 (0.016)	1.320 (0.000)	0.233 (0.414)	-0.134 (0.583)
Log(Pre-Period Patents)	0.024 (0.611)	-0.174 (0.204)	-0.095 (0.013)	0.767 (0.000)	0.697 (0.000)
Additional Pre-Period Characteristics	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y
Observations	1,165	1,165	1,165	1,165	1,165
N. of Firms	106	106	106	106	106

P-values in parentheses, based on robust standard errors clustered by firm.

5.3 Heterogeneous Treatment Effects: Pro-Patent Management

In this section, we explore the potential for heterogeneity in our treatment effects. Specifically, we interact our difference-in-differences measures with a new variable: the log-transformed number of pre-period patent filings for each of our firms. High-patent firms are considered to have pro-patent management practices (Bhaskarabhatla and Hegde 2014). The results for pro-patent management firms indicate that there is a strong tendency to diversify one's IP portfolio following the piracy shock: firms with more pre-period patent filings tend to differentially increase their copyright and trademark filings more

than firms with fewer pre-period patent filings. The effect sizes are quite large for these measures: a firm with twice as many pre-period patents can expect to have more than double the piracy-copyright elasticity of a “baseline” firm, and an approximately 50%-larger piracy-trademark elasticity. There is no difference across this margin for R&D spending or patenting. Both high-patent and low-patent firms exhibit similar increases in these measures following the rise in software piracy. Our results suggest that pro-patent-management firms are leaning into higher levels of innovative activity and diversifying their IP portfolios.

Table 6: Heterogeneous Treatment Effects - Interaction with Pre-Period Patents

	(1) 2SLS DV = Log(R&D)	(2) IV-Poisson DV = Copyright Count	(3) IV-Poisson DV = Trademark Count	(4) IV-Poisson DV = Patent Count	(5) IV-Poisson DV = Cite- Weighted Patent Count
VARIABLES					
<u>Instrumented Variables:</u>					
Log(Times Pirated) X Window (2001-2002) X Log(Pre-Period Patents)	-0.077 (0.166)	0.201 (0.025)	0.069 (0.221)	-0.002 (0.934)	-0.037 (0.288)
Log(Times Pirated) X Post-Long (2003-2007) X Log(Pre-Period Patents)	-0.079 (0.210)	0.176 (0.025)	0.087 (0.005)	0.055 (0.267)	-0.044 (0.287)
Log(Times Pirated) X Window (2001-2002)	0.358 (0.069)	0.316 (0.002)	-0.251 (0.364)	0.092 (0.425)	0.201 (0.127)
Log(Times Pirated) X Post (2003-2007)	0.480 (0.076)	0.000 (0.999)	0.027 (0.839)	0.082 (0.498)	0.264 (0.014)
Log(Total Times Pirated) X Log(Pre-Period Patents)	0.058 (0.056)	-0.023 (0.536)	-0.013 (0.377)	-0.177 (0.000)	-0.148 (0.002)
Log(Total Times Pirated)	-0.242 (0.080)	0.046 (0.647)	-0.021 (0.742)	0.314 (0.001)	0.348 (0.006)
<u>Controls:</u>					
Log(Firm Age)	Y	Y	Y	Y	Y
Pre-Period Characteristics	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y
Observations	1,165	1,165	1,165	1,165	1,165
N. of Firms	106	106	106	106	106

P-values in parentheses, based on robust standard errors clustered by firm.

Taking stock of our triple-interaction analysis, our findings support Hypothesis 3. Rather than exhibiting a greater response across all measures of innovative activity, high-patent firms only “outperform” other treated firms in the areas of copyrights and trademarks, diversifying their IP portfolio. These results suggest the potential for significant complementarity between patents and other forms of intellectual property, especially in the context of a dynamic response to an unexpected rise in product-market imitation.

5.4 Robustness Tests

We perform a range of robustness tests to further explore the details of our main findings. We first examine the pattern of piracy in our sample, confirming that pre-period at-risk status is a strong predictor of binary measures of firm-level piracy, in addition to the continuous measures of piracy presented in Table 3. Similarly, we show that using piracy rates directly, without a two-stage instrumental-variables methodology, generates similar coefficients to those we obtain in Table 4.²⁰ We replicate our primary analysis in both narrower samples (e.g., without Microsoft, and with the subset of firms with annual revenues of at least \$500M in the pre-period) and in broader samples (e.g., the full, unbalanced sample described in Table 2). We also explore alternate specifications, including a shorter pre-period, fewer control variables, and outcome variables targeting IP related to computers and software.²¹ When adding an additional post-piracy period lasting through 2013, the impact of piracy seems to be quite durable for all outcomes except raw patent counts, lasting over a decade beyond the initial piracy shock. Across a variety of robustness measures, we find the same broad pattern of results depicted in our baseline analysis in Table 4.

In addition to the sensitivity analysis above, we follow established practices in the difference-in-differences literature (Angrist and Pischke 2009, Lechner 2010) and explicitly test for a difference in pre-period (linear) time trends between our treatment and control groups.²² In all cases, we find no significant effects in the pre-piracy period, suggesting that the parallel trends assumption holds across all outcome measures for our sample of firms.

6. DISCUSSION AND CONCLUSIONS

A tenet of management science is that large, established firms are slow to respond to change, and, as a result, organizations often suffer poor financial performance and even death (Christensen 2013, Henderson and Clark 1990, Rumelt 1995, Tripsas and Gavetti 2000). However, this paper shows that, on

²⁰ Appendix Table A4 presents the direct, non-instrumented impact of piracy on innovative activity.

²¹ Appendix Table A5 depicts the analysis of computer-related IP outcomes and presents the results using at-risk hardware firms as a control group.

²² Appendix Table A3 displays the results of our parallel-trends test and our placebo-shock analysis.

average, large, public firms in the software industry were relatively quick to respond to changes in their IP environment due to digital piracy, particularly those firms with established pro-patent management practices. This supports previous work encouraging open innovation practices, which focuses primarily on the disclosure and licensing of patents, to increase organizational profitability and the diffusion of innovation (Chesbrough 2003). We extend this theory to include copyrights and trademarks to show the extent to which formal IP rights are interdependent and complementary: IP reinforces one other, and this interplay is a powerful strategic tool for firms in the software industry.

This paper also contributes to our understanding of the dynamic nature of firms' management of their IP portfolio. R&D spending, copyrights, trademarks, and patents are each indicators of innovative activity, with unique strengths and weaknesses in terms of their effectiveness to capture value for a firm. Unlike traditional goods, digital products can be distributed at near-zero marginal costs; therefore, firms' responses in how they manage this IP when faced with increased competition in the digital economy may differ from firms' responses to competition in traditional markets. This paper finds that at-risk software firms react strongly to the threat of piracy with a consistent upward trend of R&D spending and applications for copyrights, trademarks, and patents in the post-BitTorrent economy. The rise in innovation supports the interpretation of piracy as a form of product-market competition, as prior work has documented a positive relationship between competition and innovation in contexts where firms enjoy significant market power (Aghion et al. 2005, Grahovac and Miller 2009).

Looking more closely at effect sizes, our estimated elasticities of the impact of piracy on R&D spending, copyrights, trademarks, and patents range from approximately 0.05 to 0.30 depending on outcome measure and specification. At first glance, these might seem like relatively inelastic responses to fluctuations in piracy rates: for example, a 10% increase in piracy leads to a 2.8% increase in R&D spending, or an average of \$3M per year per firm. Our study is therefore consistent with previous literature that finds low-to-moderate impacts of piracy on innovation in other contexts. At the same time, it is important to

remember that piracy rates are highly volatile. As a result, even under these estimated elasticities, our results represent substantial, long-lasting impacts to firm strategy and IP management practices among the large, public firms in our sample.

6.1 Limitations and Generalizability

This paper has broad applicability to contexts where incumbent firms face product-market imitation and similar challenges. However, some of our results should be interpreted with caution. Specifically, in the analysis of the long-term effects of piracy, the number of firms and thus the number of observations decline, as firms disappear from the dataset due to mergers, acquisitions, or closure. While this trend is present for both treatment and control firms, it is important to interpret our findings as being conditional on firm survival. Relatedly, we do not explore the impact of piracy on rates of entry into the software industry or on the behavior of small or privately held firms. In addition, it is impossible to determine what invention or product is being protected by which IP instrument; instead, we bundle innovation at the firm-level. Finally, it is important to emphasize that our measures of innovative activity are inherently imperfect: trade secrets are generally unobservable, R&D expenditures may have highly uncertain outcomes, and IP filing dates may have significant lags from the timing of the original innovation. We believe that future work may be able to address some of these limitations, while others are intrinsic to the fields of strategy and innovation.

The extent to which firms can generate and exploit economically useful knowledge and rely on IP to protect information goods against spillovers, imitation, or theft, has a profound impact on firm strategy (W. M. Cohen et al. 2000, Rumelt 1984, Teece 1986). When extrapolating our conclusions to other contexts, we expect external validity to be strongest for platform industries dominated by older firms that rely on network effects and large, existing portfolios of IP. These industries include 3D printing, artificial intelligence, and video games. Furthermore, we expect our findings to be most applicable to cases where pirated products act as substitutes to legal product releases, rather than cases of product complementarity, such as movie sequels. As the economy transitions toward being increasingly technology-driven and knowledge-

based, more value will be created in markets where digital piracy is possible. Globalization and specialization have drastically decreased the price for the hardware and infrastructure required to support software products. Software piracy is likely to be an increasingly important competitive threat to firms' value-capture strategies for information goods as these trends continue in the foreseeable future.

6.2 Managerial Implications

In recent years, partly as a response to piracy, firms have embraced subscription-based models. However, competition between subscription-based business models, especially those offering exclusive content, has sometimes decreased access to popular products and increased consumers' level of "subscription fatigue" (Newman et al. 2019). Additionally, developing nations have limited access to the required Internet infrastructure to sustain subscription-based services, and government policy has been known to ban certain software products entirely (see Venezuelan (Cohen and Ellsworth 2019) or China policy). This may help explain why digital piracy via BitTorrent is once again on the rise (Bode 2018).

Given the consistent threat of software piracy, managers may need new strategies to appropriate returns from innovation. Our findings suggest that obtaining different forms of IP over varying time horizons may well generate synergies. Firms in our sample exhibited a strong rise in innovative activity following the piracy shock, but they also shifted their IP strategies to pursue a more diversified portfolio through increased filings of copyrights and trademarks. The early rise in copyrights and trademarks in the immediate aftermath of the piracy shock is consistent with the fundamental characteristics of these instruments. In addition, diversification reflects the importance of complementarities among different IP instruments, and it implies that shoring up the gaps in an IP portfolio may be more critical to firms than building on existing strengths. Note that we did not find similar diversification patterns for firms who emphasize either copyrights or trademarks in the pre-period. Copyrights and trademarks are not likely to perform as perfect substitutes for

patents in terms of the benefits they provide, such as licensing and asset valuation.²³ This may help explain why starting with pro-patent management practices first is particularly effective in positioning firms to capture value from a broad spectrum of IP instruments later.

Whether through legal action or authentication keys, attempts to eliminate piracy have been historically ineffective. While firm performance is outside the scope of this study, initial estimates show firms in our sample experienced significant increases in both revenue and profit after the piracy shock. Our findings suggest that strategies encouraging new product development and a continuous commitment to innovation may offset the negative effects of digital piracy in certain industries.

REFERENCES

- Ackermann, K., Bradley, W. A., and Cameron, J. (2020) Who profits from pirates? Evidence from the movie industry, *Working Paper*.
- Aghion P, Bloom N, Blundell R, Griffith R, Howitt P (2005) Competition and Innovation: an Inverted-U Relationship. *The Quarterly Journal of Economics* 120(2):701–728.
- Alliance BS *Ninth Annual BSA Global Software Piracy Study*
- Amara N, Landry R, Traoré N (2008) Managing the protection of innovations in knowledge-intensive business services. *Research Policy* 37(9):1530–1547.
- Angrist J, Pischke JS (2014) *Mastering 'Metrics: The Path from Cause to Effect* 1st ed. (Princeton University Press).
- Angrist JD, Pischke JS (2009) *Mostly harmless econometrics: an empiricist's companion* (Princeton University Press, Princeton).
- Arora A (1995) Licensing Tacit Knowledge: Intellectual Property Rights And The Market For Know-How. *Economics of Innovation and New Technology* 4(1):41–60.
- Arora A (1997) Patents, licensing, and market structure in the chemical industry. *Research policy* 26(4–5):391–403.
- Arora A, Ceccagnoli M (2006) Patent Protection, Complementary Assets, and Firms' Incentives for Technology Licensing. *Management Science* 52(2):293–308.
- Arundel A (2001) The relative effectiveness of patents and secrecy for appropriation. *Research policy* 30(4):611–624.
- Arundel A, Kabla I (1998) What percentage of innovations are patented? empirical estimates for European firms. *Research Policy* 27(2):127–141.
- Athey S, Stern S (2015) The Nature and Incidence of Software Piracy: Evidence from Windows. *Economic Analysis of the Digital Economy*. (National Bureau of Economic Research, Inc), 443–477.
- Awokuse TO, Yin H (2010) Intellectual property rights protection and the surge in FDI in China. *Journal of Comparative Economics* 38(2):217–224.

²³ Moving beyond the timeframe of our sample period, we would expect that copyrights and trademarks would gain in importance as instruments to protect and appropriate value from innovation in the wake of the 2014 *Alice Corp. v. CLS Bank International* U.S. Supreme Court decision, which introduced strict limits on the patentability of software.

- Bai J, Waldfogel J (2012) Movie piracy and sales displacement in two samples of Chinese consumers. *Information Economics and Policy* 24(3–4):187–196.
- Belleflamme P, Peitz M (2010) *Digital Piracy: Theory* (CESifo Group Munich).
- Berman B (2019) \$21 trillion in U.S. intangible assets is 84% of S&P 500 value – IP rights and reputation included.
- Bertrand M, Black S, Jensen S, Lleras-Muney A (2019) Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway. *Review of Economic Studies* 86(1):191–239.
- Besen SM (1986) Private copying, reproduction costs, and the supply of intellectual property. *Information Economics and Policy* 2(1):5–22.
- Besen SM, Raskind LJ (1991) An introduction to the law and economics of intellectual property. *Journal of economic perspectives* 5(1):3–27.
- Bessen J, Hunt RM (2007) An Empirical Look at Software Patents. *J Economics Management Strategy* 16(1).
- Bhaskarabhatla A, Hegde D (2014) An organizational perspective on patenting and open innovation. *Organization Science* 25(6):1744–1763.
- Bhattacharjee S, Gopal RD, Lertwachara K, Marsden JR, Telang R (2007) The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts. *Management Science* 53(9):1359–1374.
- Blind K, Edler J, Frietsch R, Schmoch U (2006) Motives to patent: Empirical evidence from Germany. *Research Policy* 35(5):655–672.
- Bode K (2018) The Rise of Netflix Competitors Has Pushed Consumers Back Toward Piracy. *Vice*.
- Campbell-Kelly M (2004) Not All Bad: An Historical Perspective on Software Patents. *Mich. Telecomm. & Tech. L. Rev.* 11:191.
- Carrier MA (2010) The Pirate Bay, Grokster, and Google. *Journal of Intellectual Property Rights* 15:7–18.
- Charny B (2003) Nokia confirms N-Gage hack, pursues pirates. *CNET*. Retrieved (June 24, 2021), <https://www.cnet.com/news/nokia-confirms-n-gage-hack-pursues-pirates/>.
- Chesbrough HW (2003) *Open innovation: The new imperative for creating and profiting from technology* (Harvard Business Press).
- Christensen CM (2013) *The innovator's dilemma: when new technologies cause great firms to fail* (Harvard Business Review Press).
- Cockburn IM, MacGarvie MJ (2011) Entry and Patenting in the Software Industry. *Management Science* 57(5):915–933.
- Cohen L, Ellsworth B (2019) Venezuela designers turn to piracy after Adobe announces it will cut service. *Reuters* (October 9) <https://www.reuters.com/article/us-venezuela-politics-adobe-idUSKBN1WN25L>
- Cohen W, Nelson R, Walsh J (2000) *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* (National Bureau of Economic Research, Cambridge, MA).
- Conner KR, Rumelt RP (1991) Software Piracy: An Analysis of Protection Strategies. *Management Science* 37(2):125–139.
- Costello K, Rimol M Gartner Forecasts Worldwide IT Spending to Reach \$4 Trillion in 2021. *Gartner Press Release*. Retrieved (July 23, 2021), <https://www.gartner.com/en/newsroom/press-releases/2021-04-07-gartner-forecasts-worldwide-it-spending-to-reach-4-trillion-in-2021>.
- Danaher B, Smith MD (2014) Gone in 60 Seconds: The Impact of the Megaupload Shutdown on Movie Sales. *International Journal of Industrial Organization* 33:1–8.
- Danaher B, Smith MD, Telang R (2014) Piracy and Copyright Enforcement Mechanisms. *Innovation Policy and the Economy* 14:25–61.

- Danaher B, Smith MD, Telang R, Chen S (2014) The Effect of Graduated Response Anti-Piracy Laws on Music Sales: Evidence from an Event Study in France: Evidence from an Event Study in France. *The Journal of Industrial Economics* 62(3):541–553.
- D'Aveni RA, Gunther RE (1994) *Hypercompetition: managing the dynamics of strategic maneuvering* (The Free Press; Maxwell Macmillan Canada; Maxwell Macmillan International, New York: Toronto: New York).
- De Vany AS, Walls WD (2007) Estimating the Effects of Movie Piracy on Box-office Revenue. *Review of Industrial Organization* 30(4):291–301.
- Edvardsson B, Kristensson P, Magnusson P, Matthing J, Gustafsson A (2006) *Involving Customers In New Service Development* (World Scientific Publishing Company).
- Friedman DD, Landes WM, Posner RA (1991) Some economics of trade secret law. *Journal of Economic Perspectives* 5(1):61–72.
- Givon M, Mahajan V, Muller E (1995) Software piracy: Estimation of lost sales and the impact on software diffusion. *Journal of Marketing* 59(1):29–37.
- Goldfarb A, Greenstein SM, Tucker C eds. (2015) *Economic analysis of the digital economy* (The University of Chicago Press, Chicago; London).
- Gotsch M, Hipp C (2014) Using Trademarks to Measure Innovation in Knowledge-Intensive Business Services. *Technology Innovation Management Review* 4(5):18–30.
- Graham S, Mowery D (2006) The use of intellectual property in software: implications for open innovation.
- Graham S, Vishnubakat S (2013) Of Smart Phone Wars and Software Patents. *Journal of Economic Perspectives* 27(1):67–86.
- Grahovac J, Miller DJ (2009) Competitive advantage and performance: the impact of value creation and costliness of imitation. *Strat. Mgmt. J.* 30(11):1192–1212.
- Greenstein S, Lerner J, Stern S (2013) Digitization, innovation, and copyright: What is the agenda? *Strategic Organization* 11(1):110–121.
- Hall B, Helmers C, Rogers M, Sena V (2014) The Choice between Formal and Informal Intellectual Property: A Review. *Journal of Economic Literature* 52(2):1–50.
- Hall BH (2004) Exploring the patent explosion. *The Journal of Technology Transfer* 30(1–2):35–48.
- Hall BH, MacGarvie M (2010) The private value of software patents. *Research Policy* 39(7):994–1009.
- Hall BH, Ziedonis RH (2001) The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979–1995. *The RAND Journal of Economics* 32(1):101.
- Harabi N (1995) Appropriability of technical innovations an empirical analysis. *Research Policy* 24(6):981–992.
- Henderson RM, Clark KB (1990) Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative science quarterly*:9–30.
- Horstmann I, MacDonald GM, Slivinski A (1985) Patents as information transfer mechanisms: To patent or (maybe) not to patent. *Journal of Political Economy* 93(5):837–858.
- Hurmelinna-Laukkanen P, Puumalainen K (2007) Nature and dynamics of appropriability: strategies for appropriating returns on innovation. *R & D Management* 37(2):95–112.
- Iacus SM, King G, Porro G (2012) Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis* 20(1):1–24.
- Jaffe AB, Lerner J, Stern S, National Bureau of Economic Research, Innovation Policy and the Economy Group (2001) *Innovation policy and the economy* (MIT Press, Cambridge, Mass.).

- Jaisingh J (2009) Impact of piracy on innovation at software firms and implications for piracy policy. *Decision Support Systems* 46(4):763–773.
- Kariithi NK (2011) Is the Devil in the Data? A Literature Review of Piracy Around the World: Literature Review of Piracy Around the World. *The Journal of World Intellectual Property* 14(2):133–154.
- Kortum S, Lerner J (1999) What is behind the recent surge in patenting? *Research policy* 28(1):1–22.
- Kretschmer T, Peukert C (2014) Video killed the radio star? Online music videos and digital music sales. *CEP Discussion Papers* (1265).
- Lahiri A, Dey D (2013) Effects of Piracy on Quality of Information Goods. *Management Science* 59(1):245–264.
- Landes WM, Posner RA (1989) An Economic Analysis of Copyright Law. *The Journal of Legal Studies* 18(2):325–363.
- Laursen K, Salter A (2006) Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal* 27(2):131–150.
- Lechner M (2010) The Estimation of Causal Effects by Difference-in-Difference Methods Estimation of Spatial Panels. *Foundations and Trends® in Econometrics* 4(3):165–224.
- Lee JF (2018) Purchase, pirate, publicize: Private-network music sharing and market album sales. *Information Economics and Policy* 42:35–55.
- Lemley MA, Cohen JE (2000) Patent Scope and Innovation in the Software Industry. *SSRN Journal*.
- Levin RC, Klevorick AK, Nelson R, Winter S (1987) Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity* 18(3, Special Issue on Microeconomics):783–832.
- Liebowitz SJ (1985) Copying and Indirect Appropriability: Photocopying of Journals. *Journal of Political Economy* 93(5):945–957.
- Lu S, Wang X (Shane), Bendle N (2020) Does Piracy Create Online Word of Mouth? An Empirical Analysis in the Movie Industry. *Management Science* 66(5):2140–2162.
- Luo H, Mortimer JH (2017) Copyright Enforcement Evidence from Two Field Experiments: Copyright Enforcement. *Journal of Economics & Management Strategy* 26(2):499–528.
- Ma L, Montgomery AL, Singh PV, Smith MD (2014) An Empirical Analysis of the Impact of Pre-Release Movie Piracy on Box Office Revenue. *Information Systems Research* 25(3):590–603.
- Mansfield E (1986) Patents and Innovation: An Empirical Study. *Management Science* 32(2):173–181.
- Mendonça S, Pereira TS, Godinho MM (2004) Trademarks as an indicator of innovation and industrial change. *Research Policy* 33(9):1385–1404.
- Miller JP (2018) *Why The Pirate Bay Remains At Sea While Napster 1.0 Now Rests In Davy Jones's Locker*.
- Miric M, Jeppesen LB (2020) Does Piracy Lead to Product Abandonment or Stimulate New Product Development?: Evidence from Mobile PLATFORM-BASED Developer Firms. *Strategic Management Journal*.
- Morris S (2018) Why BitTorrent Mattered — Bittorrent Lessons for Crypto (1 of 4). *Medium*.
- Newman N, Fletcher R, Kalogeropoulos A, Nielsen RK (2019) Reuters Institute Digital News Report 2019. :156.
- Paallysaho S, Kuusisto J (2008) Intellectual property protection as a key driver of service innovation: an analysis of innovative KIBS businesses in Finland and the UK. *IJSTM* 9(3/4):268.
- Peitz M, Waelbroeck P (2006) Piracy of digital products: A critical review of the theoretical literature. *Information Economics and Policy* 18(4):449–476.
- Peukert C, Claussen J, Kretschmer T (2017) Piracy and box office movie revenues: Evidence from Megaupload. *International Journal of Industrial Organization* 52:188–215.
- Posen HE, Martignoni D (2018) Revisiting the imitation assumption: Why imitation may increase, rather than decrease, performance heterogeneity. *Strat Mgmt J* 39(5):1350–1369.

- Reavis Conner K, Rumelt RP (1991) Software Piracy: An Analysis of Protection Strategies. *Management Science* 37(2):125–139.
- Reitzig M (2004) Strategic Management of Intellectual Property. *MIT Sloan Management Review* 45(3):35–40.
- Rob R, Waldfogel J (2006) Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *The Journal of Law and Economics* 49(1):29–62.
- Roberts PW (1999) Product innovation, product-market competition and persistent profitability in the U.S. pharmaceutical industry. *Strategic Management Journal* 20(7):655–670.
- Rumelt RP (1984) Towards a strategic theory of the firm. *Competitive strategic management* 26(3):556–570.
- Rumelt RP (1995) Inertia and transformation. *Resource-based and evolutionary theories of the firm: Towards a synthesis*. (Springer), 101–132.
- Samuelson P (1984) CONTU revisited: the case against copyright protection for computer programs in machine-readable form. *Duke Lj*:663.
- Samuelson P (2010) The Uneasy Case for Software Copyrights Revisited. *Geo. Wash. L. Rev.* 79:1746.
- Schumpeter JA (1939) *Capitalism, Socialism and Democracy* (McGraw-Hill, New York).
- Shy O, Thisse JF (1999) A Strategic Approach to Software Protection*. *Journal of Economics & Management Strategy* 8:163–190.
- Somaya D, Graham SJH (2006) Vermeers and Rembrandts in the Same Attic: Complementarity between Copyright and Trademark Leveraging Strategies in Software. *SSRN Journal*.
- Stonefield S The 10 Most Valuable Trademarks. *Forbes*. Retrieved (July 23, 2021), <https://www.forbes.com/sites/seanstonefield/2011/06/15/the-10-most-valuable-trademarks/>.
- Teece DJ (1986) Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15(6):285–305.
- Tripsas M, Gavetti G (2000) Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic management journal* 21(10–11):1147–1161.
- Varian HR (2003) Buying, Sharing and Renting Information Goods. *The Journal of Industrial Economics* 48(4):473–488.
- Waldfogel J (2012) Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster. *The Journal of Law and Economics* 55(4):715–740.
- Wang Q, Li B, Singh PV (2018) Copycats vs. Original Mobile Apps: A Machine Learning Copycat-Detection Method and Empirical Analysis. *Information Systems Research* 29(2):273–291.
- Wang RD, Shaver JM (2014) Competition-driven repositioning: Competition-Driven Repositioning. *Strategic Management Journal* 35(11):1585–1604.
- Yoon C (2012) Digital piracy intention: a comparison of theoretical models. *Behaviour & Information Technology* 31(6):565–576.
- Zhao M (2006) Conducting R&D in Countries with Weak Intellectual Property Rights Protection. *Management Science* 52(8):1185–1199.
- Ziedonis RH (2004) Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms. *Management Science* 50(6):804–820.

APPENDIX

Empirical Setting

The BitTorrent communication protocol was a disruptive technology that facilitated unprecedented levels of piracy across a variety of industries in the digital economy. The majority of empirical research examining piracy focuses on media, including music and movies, where pirated copies cannot be automatically seen as, and are likely not, perfect product substitutes (Bai and Waldfogel 2012, Danaher et al. 2010, Danaher, Smith, Telang, et al. 2014, Danaher and Smith 2014, De Vany and Walls 2007, Oberholzer-Gee and Strumpf 2007, Rob and Waldfogel 2006, Waldfogel 2012a). Unlike pirated movies or music, pirated software contains identical features, quality, and performance capabilities to the original, and there is reason to believe this may invoke different competitive responses by firms. Specifically, firms must survive the sudden appearance of identical, free rival versions of their products; further, they must do so without access to price as a mechanism for market repositioning to sustain their competitive advantage and under conditions where new rivals reappear continually over time (Argyres et al. 2019, Porter and Rivkin 2000, Wang and Shaver 2014).

Unlike older systems of file sharing,²⁴ BitTorrent's open, decentralized file sharing allowed individuals to upload and download pieces of a single file concurrently, without waiting for any single user to possess the entire file. This radically reduced the time it takes to share a file. In *Fortune Magazine's* words: "Before BitTorrent, large file transfers basically operated like the world's slowest Blockbuster," (Roth and

²⁴ Software distribution has evolved continuously over the industry's history. However, the rise of digital technologies, such as centralized and de-centralized servers, cloud systems, and BitTorrent, have changed the nature of software piracy and dramatically lowered the costs to both supplying and acquiring pirated software over time. Traditional P2P file sharing, such as Napster or LimeWire, popularized the pirating of small files and mp3s by using a server to link two individuals who wanted to share files. A person would need to download and upload a file in its entirety before that file could be shared with other users.

Ryan 2005).²⁵ BitTorrent was introduced in 2001 when, without warning or announcement, BitTorrent's creator Bram Cohen first published his source code online as Linux software completely free to users (Thompson 2005). BitTorrent was initially designed for legal uses; indeed, entities including Facebook, Twitter, most Linux-based operating system distributions, and the United Kingdom have all used the protocol to distribute large data files (Hoffman 2013). As a testament to its legal validity, the BitTorrent protocol has never been shut down, nor has there been any government-driven attempt to shut it down.²⁶ Since its launch, tens of millions of people have downloaded BitTorrent clients. As of 2013, BitTorrent alone had 15–27 million concurrent users at any time (Wang and Kangasharju 2013). According to ComScore data, by June 2013, the BitTorrent Network was ranked 48 among the most visited websites in the U.S., barely trailing Netflix.com, ranked 35, and roughly the same as Time Warner, ranked 47. BitTorrent is the largest single protocol, substantially larger than HTTP or FTP (Johnsen et al. 2005). Though not intended by its designer, one of the most notable consequences of the BitTorrent protocol was a sharp and sudden increase in the piracy of large file formats, such as movies, games, and software, leading to IP infringement at unprecedented scales (Carrier 2010). The availability of BitTorrent data from The Pirate Bay has recently increased interest in this form of piracy within empirical work in the organization and management sciences (Lu, Wang and Bendle, 2020; Ackermann, Bradley and Cameron, 2020).

²⁵ In traditional P2P file sharing, a 120MB file copied to three different people by one user would take approximately two hours, maxing out the user's 3MB upload bandwidth and using 25% of each acquiring individual's download bandwidth. With BitTorrent, the same file can be shared among the three users in half the time by maxing out upload bandwidth (at a more efficient 75% downloading and uploading bandwidth by all individuals).

²⁶ For more about "Why BitTorrent Mattered," see Simon Morris' *Medium* article (parentheticals added): "Certainly there were plenty of ISP attempts to interfere with what was considered excessive BitTorrent traffic (due to piracy), but a concerted and state-sanctioned attempt to kill it never materialized. In large part this is because no-one wanted to set a legal precedent of censoring general purpose technology simply because of one set of legally problematic uses" (Morris 2018).

Figure A1: Pre-Period IP Composition for Treatment Firms

This figure presents a simplex plot of the composition of pre-period IP portfolios for each of our treatment firms. We calculate percentages after normalizing each IP counter by its in-sample median in order to highlight the relative tendencies of firms to employ each of the three IP instruments we track. The figure illustrates the significant variation in IP portfolios in our sample.

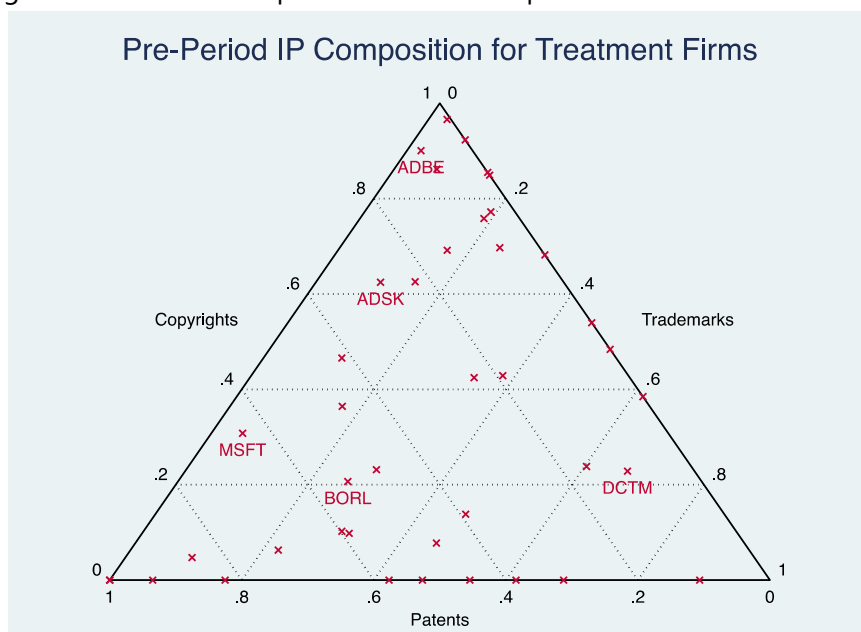


Figure A2: Treatment Effect on Citation-Weighted Patents

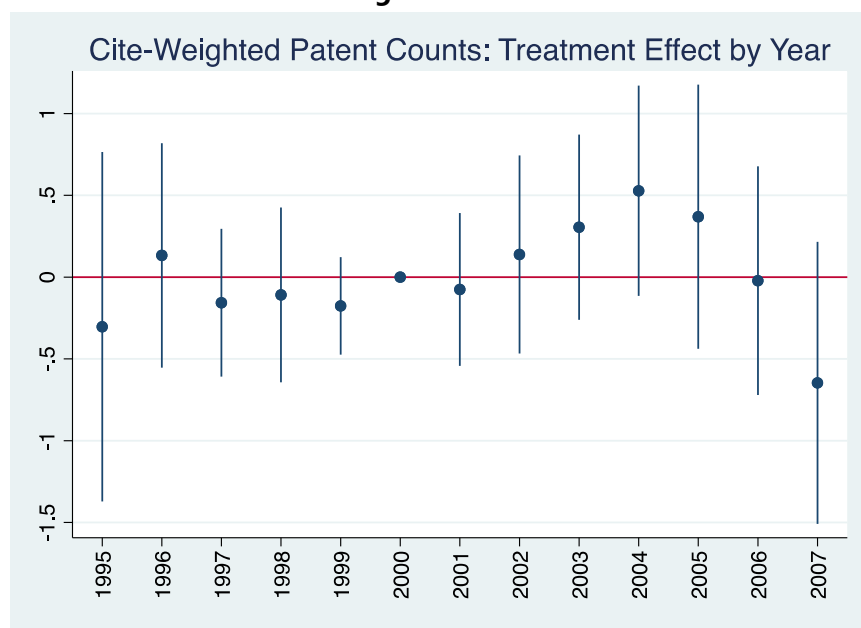


Table A1: Sample of Firms

Treatment Group: SEC Piracy-Mentioning Firms (N = 41)

<u>gvkey</u>	<u>tic</u>	<u>conm</u>
112624	BIRT	ACTUATE CORP
12540	ADBE	ADOBE INC
121493	ARBA	ARIBA INC
122159	ARTG	ART TECHNOLOGY GROUP INC
12587	ASCL	ASCENTIAL SOFTWARE CORP
1878	ADSK	AUTODESK INC
64606	BEAS	BEA SYSTEMS INC
14268	BORL	BORLAND SOFTWARE CORP
110160	BRIO	BRIO SOFTWARE INC
28822	CPTV	CAPTIVA SOFTWARE CORP
63185	CHKP	CHECK POINT SOFTWARE TECHN
62005	DCTM	DOCUMENTUM INC
113609	ENTU	ENTRUST INC
138486	EXEE	EXE TECHNOLOGIES INC
23777	FSTP	FRONTSTEP INC
60901	HNCS	HNC SOFTWARE INC
61498	HYSL	HYPERION SOLUTIONS CORP
120134	INFA	INFORMATICA CORP
106900	ISSX	INTERNET SECURITY SYSTEMS
124678	IWOV	INTERWOVEN INC
27928	INTU	INTUIT INC
60969	LGTO	LEGATO SYSTEMS INC
29381	MACR	MACROMEDIA INC
28744	MANU.1	MANUGISTICS GROUP INC
25783	MFE	MCAFEЕ INC
102690	MRNTY.1	MERANT PLC
12141	MSFT	MICROSOFT CORP
111534	MSTR	MICROSTRATEGY INC
28881	NETM	NETMANAGE INC
118307	ONXS	ONYX SOFTWARE CORP
64597	PRGN.1	PEREGRINE SYSTEMS INC
65226	QADA	QAD INC
122921	QSFT	QUEST SOFTWARE INC
65899	RNWK	REALNETWORKS INC
31622	RMDY	REMEDY CORP
63180	SEBL	SIEBEL SYSTEMS INC
28758	SPSS	SPSS INC
123998	SWKH	SWK HOLDINGS CORP
15855	SYMC	SYMANTEC CORP
29356	VRTS.1	VERITAS SOFTWARE CORP
118445	VIGN	VIGNETTE CORP

Control Group: Non-Piracy-Mentioning Firms (N = 65)

<u>gvkey</u>	<u>tic</u>	<u>conm</u>
61656	ACCL	ACCELRYIS INC
31564	ACIW	ACI WORLDWIDE INC
61726	ADEP	ADEPT TECHNOLOGY INC
61557	ADVS	ADVENT SOFTWARE INC
117701	ALLR.1	ALLAIRE CORP
65072	AMCS	AMICAS INC
30870	AZPN	ASPEN TECHNOLOGY INC
60797	AVNT	AVANT CORP
30688	BOBJY	BUSINESS OBJECTS SA
3310	CA	CA INC
31143	CAPA.	CAPTARIS INC
110721	CANI	CARREKER CORP
61355	CKFR	CHECKFREE CORP
61676	CTXS	CITRIX SYSTEMS INC
14459	COGN	COGNOS INC
110766	CMTOQ	COM21 INC
26011	CPWR.1	COMPUWARE CORP
17080	CREL	COREL CORP
31660	DSTM	DATASTREAM SYSTEMS INC
64423	PROJ	DELTEK INC
60950	DRTE	DENDRITE INTERNATIONAL INC
123995	EPNY	E.PIPHANY INC
65466	JDEC	EDWARDS J D & CO
25774	EFII	ELECTRONICS FOR IMAGING INC
25091	3ELVN.	ELEVON INC
122137	ENGA.	ENGAGE INC
25859	EPIC	EPICOR SOFTWARE CORP -OLD
25631	EPRE	EPRESENCE INC
13525	FILE	FILENET CORP
61566	IDXC	IDX SYSTEMS CORP
62285	IINT	INDUS INTERNATIONAL INC
63648	3IMIC	INDUSRI-MATEMATIK INTL CORP
120839	INET.1	INET TECHNOLOGIES INC
61573	INFM	INFINIUM SOFTWARE INC
111488	INKT	INKTOMI CORP
120845	IACT	INTERACT COMMERCE CORP
25302	INTZ	INTRUSION INC
62399	JDAS	JDA SOFTWARE GROUP INC
23034	MAPX.1	MAPICS INC
65031	MCTR	MERCATOR SOFTWARE INC
29095	MERQ	MERCURY INTERACTIVE CORP

Control Group, Continued

<u>gvkey</u>	<u>tic</u>	<u>conm</u>
126597	MSLV	METASOLV INC
66468	MUSE	MICROMUSE INC
30094	MROI	MRO SOFTWARE INC
31607	NATI	NATIONAL INSTRUMENTS CORP
64927	NEON.	NEW ERA OF NETWORKS INC
14626	PTEC	PHOENIX TECHNOLOGIES LTD
120300	PRSF.	PORTAL SOFTWARE INC
24352	PRGS	PROGRESS SOFTWARE CORP
18699	PTC	PTC INC
63758	QDHC	QUADRAMED CORP
11135	RATL	RATIONAL SOFTWARE CORP
122841	RHT	RED HAT INC
134452	SBYN	SEEBEYOND TECHNOLOGY CORP
118321	SRNA.1	SERENA SOFTWARE INC
13902	SDRC.1	STRUCTURAL DYNAMICS RESEARCH
24409	SY.3	SYBASE INC
24975	SNPS	SYNOPSIS INC
28325	3TTLA	TARANTELLA INC
122061	TIBX	TIBCO SOFTWARE INC
66368	VRSN	VERISIGN INC
61401	VRTY.	VERITY INC
24489	VRCC	VERTICAL COMMUNICATIONS INC
123937	VITR	VITRIA TECHNOLOGY INC
16531	ZIXI	ZIX CORP

Table A2: Correlation Matrix

#	Variable Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Log (R&D Expenditures)																								
2	Copyright Count	0.26																							
3	Trademark Count	0.44	0.28																						
4	Patent Count	0.38	0.18	0.79																					
5	Citation-Weighted Patent Count	0.42	0.22	0.87	0.95																				
6	Year	0.19	-0.03	-0.03	0.10	0.07																			
7	Log(Firm Age)	0.44	0.17	0.14	0.15	0.16	0.58																		
8	Ever Pirated	0.38	0.20	0.27	0.20	0.23	-0.09	0.22																	
9	Total Times Pirated	0.43	0.53	0.60	0.51	0.58	-0.05	0.20	0.46																
10	Pre-Period Piracy Indicator	0.19	0.11	0.15	0.13	0.15	0.04	-0.01	0.27	0.28															
11	Pre-Period Piracy 10-Ks	0.32	0.27	0.33	0.26	0.30	-0.02	0.11	0.47	0.58	0.82														
12	Pre-Period SEC 10-K Filings	0.27	0.26	0.22	0.16	0.18	-0.27	0.37	0.25	0.35	-0.06	0.22													
13	IPO Year	-0.34	-0.29	-0.26	-0.18	-0.20	0.30	-0.50	-0.39	-0.36	0.05	-0.18	-0.77												
14	Log(Pre-Period R&D)	0.72	0.35	0.50	0.39	0.44	-0.13	0.31	0.48	0.54	0.18	0.38	0.51	-0.60											
15	Log(Pre-Period Revenue)	0.66	0.34	0.49	0.39	0.45	-0.15	0.34	0.47	0.52	0.13	0.32	0.54	-0.65	0.93										
16	Log(Pre-Period Market Value)	0.56	0.23	0.41	0.33	0.39	0.12	0.05	0.27	0.39	0.29	0.31	-0.13	0.00	0.58	0.55									
17	Log(Pre-Period Assets)	0.67	0.29	0.48	0.37	0.43	-0.03	0.23	0.39	0.46	0.20	0.31	0.26	-0.39	0.85	0.83	0.83								
18	Log(Pre-Period Employees)	0.62	0.31	0.53	0.43	0.49	-0.08	0.27	0.36	0.47	0.11	0.26	0.42	-0.50	0.87	0.92	0.61	0.83							
19	Log(Pre-Period Acquisitions)	0.40	0.24	0.20	0.14	0.16	-0.10	0.28	0.25	0.22	-0.06	0.11	0.50	-0.51	0.62	0.66	0.30	0.53	0.64						
20	Average Pre-Period Net Margin	0.19	0.15	0.15	0.13	0.15	-0.18	0.22	0.22	0.22	-0.06	0.12	0.48	-0.46	0.30	0.44	-0.12	0.03	0.25	0.36					
21	Average Pre-Period Leverage	-0.07	-0.01	-0.03	-0.02	-0.03	0.01	-0.01	-0.05	-0.05	-0.15	-0.12	0.01	0.02	-0.04	-0.01	-0.14	-0.05	0.04	0.21	-0.10				
22	Average Pre-Period Revenue Growth	0.04	-0.08	-0.03	-0.01	-0.01	0.25	-0.26	-0.13	-0.08	0.19	0.03	-0.56	0.59	-0.15	-0.26	0.55	0.18	-0.11	-0.19	-0.58	-0.07			
23	Log(Pre-Period Copyrights)	0.42	0.48	0.42	0.30	0.33	-0.09	0.29	0.39	0.60	0.13	0.40	0.52	-0.54	0.62	0.65	0.31	0.52	0.63	0.46	0.29	-0.03	-0.25		
24	Log(Pre-Period Trademarks)	0.45	0.33	0.54	0.37	0.42	-0.07	0.25	0.48	0.55	0.22	0.42	0.31	-0.46	0.62	0.60	0.45	0.59	0.56	0.39	0.19	-0.04	-0.15	0.53	
25	Log(Pre-Period Patents)	0.58	0.34	0.57	0.46	0.53	-0.02	0.26	0.47	0.66	0.24	0.45	0.33	-0.40	0.73	0.69	0.61	0.70	0.62	0.35	0.19	-0.17	0.03	0.51	0.72

Table A3: Pre-Period Parallel Trends Test

This table tests for the presence of a (linear) treatment-specific time trend during the pre-period of 1991-2000. A null result here supports the assumption of parallel trends for our treatment and control firms.

VARIABLES	(1) OLS DV = Log(R&D)	(2) Poisson DV = Copyright Count	(3) Poisson DV = Trademark Count	(4) Poisson DV = Patent Count	(5) Poisson DV = Cite- Weighted Patent Count
Pre-Period Piracy 10-Ks X Year	0.007 (0.396)	-0.014 (0.108)	-0.002 (0.859)	0.012 (0.226)	0.018 (0.178)
<u>Selected Controls:</u>					
Pre-Period Piracy 10-Ks	-0.018 (0.354)	0.021 (0.558)	-0.010 (0.350)	-0.049 (0.015)	-0.016 (0.617)
Pre-Period SEC 10-K Filings	-0.056 (0.267)	0.199 (0.225)	-0.010 (0.787)	0.038 (0.580)	0.048 (0.646)
Log(Firm Age)	0.393 (0.018)	0.483 (0.048)	-0.203 (0.216)	0.265 (0.364)	0.693 (0.055)
IPO Year	0.032 (0.314)	0.034 (0.556)	-0.014 (0.590)	0.028 (0.523)	0.099 (0.102)
Log(Pre-Period R&D)	1.043 (0.000)	-0.324 (0.273)	0.126 (0.191)	-0.100 (0.476)	0.085 (0.667)
Log(Pre-Period Copyrights)	0.034 (0.258)	1.129 (0.000)	-0.010 (0.555)	0.011 (0.769)	-0.110 (0.068)
Log(Pre-Period Trademarks)	-0.063 (0.217)	0.398 (0.101)	1.412 (0.000)	-0.068 (0.552)	0.058 (0.722)
Log(Pre-Period Patents)	0.100 (0.005)	-0.223 (0.106)	-0.049 (0.081)	1.354 (0.000)	1.022 (0.000)
Additional Pre-Period Characteristics	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y
Observations	623	623	623	623	623
R-squared	0.736	0.873	0.579	0.935	0.941
N. of Firms	106	106	106	106	106

P-values in parentheses, based on robust standard errors clustered by firm.

Table A4: Piracy and Innovation

This table ignores the potential endogeneity and other concerns with our measure of observed piracy, and runs non-instrumented OLS and Poisson estimators to explore the non-identified association between software piracy and firms' innovative output.

VARIABLES	(1) OLS DV = Log(R&D)	(2) Poisson DV = Copyright Count	(3) Poisson DV = Trademark Count	(4) Poisson DV = Patent Count	(5) Poisson DV = Cite- Weighted Patent Count
Log(Times Pirated) X Post (2001-2007)	0.170 (0.025)	0.192 (0.016)	0.157 (0.000)	0.089 (0.038)	0.065 (0.076)
Log(Total Times Pirated)	-0.080 (0.080)	0.026 (0.735)	-0.020 (0.280)	0.127 (0.014)	0.150 (0.003)
<u>Selected Controls:</u>					
Pre-Period SEC 10-K Filings	-0.003 (0.948)	0.049 (0.787)	0.089 (0.160)	0.141 (0.351)	0.121 (0.417)
Log(Firm Age)	0.601 (0.000)	1.671 (0.001)	-0.004 (0.982)	0.450 (0.340)	0.209 (0.401)
IPO Year	0.070 (0.005)	0.208 (0.014)	0.031 (0.282)	0.118 (0.027)	0.091 (0.086)
Log(Pre-Period R&D)	0.971 (0.000)	-0.090 (0.767)	-0.049 (0.711)	0.061 (0.860)	0.100 (0.670)
Log(Pre-Period Copyrights)	0.064 (0.177)	0.879 (0.000)	-0.045 (0.104)	-0.198 (0.016)	-0.284 (0.000)
Log(Pre-Period Trademarks)	0.041 (0.588)	0.482 (0.014)	1.294 (0.000)	0.146 (0.556)	-0.138 (0.556)
Log(Pre-Period Patents)	0.025 (0.604)	-0.170 (0.227)	-0.116 (0.000)	0.759 (0.000)	0.689 (0.000)
Additional Pre-Period Characteristics	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y
Observations	1,165	1,165	1,165	1,165	1,165
R-squared	0.639	0.788	0.618	0.935	0.927
N. of Firms	106	106	106	106	106

P-values in parentheses, based on robust standard errors clustered by firm.

Table A5: Analysis of Computer-Specific IP

This table tracks intellectual property relevant to computers and software. Specifically, we track copyrights over "computer files," trademarks in classes 9, 35, 37, 38, and 42, and patents in NBER sub-categories 22, 23, 24, 25, and 46. For patent citations, we track all citations to computer-relevant original patents.

VARIABLES	(1) DV = relevant copyrights	(2) DV = relevant trademarks	(3) DV = relevant patents	(4) DV = citations to relevant patents
<u>Instrumented Variables:</u>				
Log(Times Pirated) X Window (2001-2002)	0.385 (0.000)	0.059 (0.344)	0.004 (0.913)	-0.012 (0.689)
Log(Times Pirated) X Post-Long (2003-2007)	0.159 (0.066)	0.203 (0.000)	0.113 (0.002)	0.093 (0.014)
Log(Total Times Pirated)	0.029 (0.807)	-0.041 (0.247)	0.077 (0.274)	0.187 (0.002)
<u>Controls:</u>				
Log(Firm Age)	Y	Y	Y	Y
Pre-Period Characteristics	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Observations	1,165	1,165	1,165	1,165
N. of Firms	106	106	106	106

IV-Poisson specifications; P-values in parentheses, based on robust standard errors clustered by firm.

Table A6: Reduced-Form Analysis

This table implements a direct difference-in-differences estimation of the impact of the post-piracy time periods on our treatment group of at-risk-of-piracy firms, based on pre-period 10-K filings that mention piracy. For each outcome variable, we include one specification using pre-period firm characteristics and another with firm-level fixed effects.

VARIABLES	(1) OLS DV = Log(R&D)	(2) FE-OLS	(3) Poisson DV = Copyright Count	(4) FE-Poisson	(5) Poisson DV = Trademark Count	(6) FE-Poisson	(7) Poisson DV = Patent Count	(8) FE-Poisson	(9) Poisson DV = Cite-Weighted Patent Count	(10) FE-Poisson
Pre-Period Piracy 10-Ks X Window (2001-2002)	0.087 (0.049)	0.098 (0.039)	0.248 (0.007)	0.246 (0.001)	0.083 (0.104)	0.074 (0.110)	0.008 (0.789)	0.002 (0.934)	-0.008 (0.757)	-0.022 (0.392)
Pre-Period Piracy 10-Ks X Post-Long (2003-2007)	0.152 (0.029)	0.168 (0.018)	0.115 (0.148)	0.129 (0.015)	0.158 (0.000)	0.149 (0.000)	0.098 (0.001)	0.097 (0.000)	0.083 (0.014)	0.071 (0.013)
<u>Selected Controls:</u>										
Pre-Period Piracy 10-Ks	-0.057 (0.066)	-	-0.042 (0.363)	-	-0.055 (0.006)	-	0.005 (0.903)	-	0.081 (0.060)	-
Log(Firm Age)	0.552 (0.000)	0.635 (0.000)	1.547 (0.002)	1.452 (0.001)	-0.026 (0.872)	-0.123 (0.458)	0.629 (0.170)	0.401 (0.319)	0.297 (0.298)	0.032 (0.902)
Pre-Period Firm Characteristics	Y	-	Y	-	Y	-	Y	-	Y	-
Firm Fixed Effects	N	Y	N	Y	N	Y	N	Y	N	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,165	1,165	1,165	1,033	1,165	1,115	1,165	939	1,165	939
N. of Firms	106	106	106	90	106	101	106	84	106	84

P-values in parentheses, based on robust standard errors clustered by firm.